

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
24 June 2004 (24.06.2004)

PCT

(10) International Publication Number
WO 2004/053077 A2

(51) International Patent Classification⁷:

C12N

(74) Agents: LICATA, Jane Massey et al.; Licata & Tyrrell
P.C., 66 E. Main Street, Marlton, NJ 08053 (US).

(21) International Application Number:

PCT/US2003/038815

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NJ, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC,
SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(30) Priority Data:

60/431,123 5 December 2002 (05.12.2002) US

(84) Designated States (*regional*): ARIPO patent (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): DI-

ADEXUS, INC. [US/US]; 343 Oyster Point Boulevard,
South San Francisco, CA 94080 (US).

Published:

— without international search report and to be republished
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): MACINA, Roberto,
A. [AR/US]; 4118 Crescendo Avenue, San Jose, CA 95136
(US). TURNER, Leah, R. [US/US]; 939 Rosette Court,
Sunnyvale, CA 94086 (US). SUN, Yongming [CN/US];
551 Shoal Circle, Redwood City, CA 94065 (US). CHEN,
Huei-Mei [US/US]; 101 Lockwood Lane, Pleasant Hill,
CA 94523 (US). RODRIGUEZ, Maria [US/US]; 570 Av-
ocet, #8109, Redwood City, CA 94065 (US).

(54) Title: COMPOSITIONS, SPLICE VARIANTS AND METHODS RELATING TO BREAST SPECIFIC GENES AND PROTEINS

(57) Abstract: The present invention relates to newly identified nucleic acid molecules and polypeptides present in normal and neoplastic breast cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions containing the nucleic acid molecules, polypeptides, antibodies, agonists and antagonists of the invention and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating breast cancer and non-cancerous disease states in breast, identifying breast tissue, monitoring and identifying and/or designing agonists and antagonists of polypeptides of the invention. The uses also include gene therapy, production of transgenic animals and cells, and production of engineered breast tissue for treatment and research.

BEST AVAILABLE COPY

WO 2004/053077 A2

1 JC09 Rec'd PCT/PTO 3 JUN 2005

**COMPOSITIONS, SPLICE VARIANTS AND METHODS
RELATING TO BREAST SPECIFIC GENES AND PROTEINS**

5

INTRODUCTION

This application claims the benefit of priority from U.S. Provisional Patent Application Serial No. 60/431,123 filed December 5, 2002 which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

10 The present invention relates to newly identified nucleic acids and polypeptides present in normal and neoplastic breast cells, including fragments, variants and derivatives of the nucleic acids and polypeptides. The present invention also relates to antibodies to the polypeptides of the invention, as well as agonists and antagonists of the polypeptides of the invention. The invention also relates to compositions comprising the nucleic acids, 15 polypeptides, antibodies, post translational modifications (PTMs), variants, derivatives, agonists and antagonists thereto and methods for the use of these compositions. These uses include identifying, diagnosing, monitoring, staging, imaging and treating breast cancer and non-cancerous disease states in breast, identifying breast tissue and monitoring and identifying and/or designing agonists and antagonists of polypeptides of the invention.

20 The uses also include gene therapy, therapeutic molecules including but not limited to antibodies or antisense molecules, production of transgenic animals and cells, and production of engineered breast tissue for treatment and research.

BACKGROUND OF THE INVENTION

Breast cancer, also referred to as mammary tumor cancer, is the second most 25 common cancer among women, accounting for a third of the cancers diagnosed in the United States. One in nine women will develop breast cancer in her lifetime and about 192,000 new cases of breast cancer are diagnosed annually with about 42,000 deaths. Bevers, *Primary Prevention of Breast Cancer*, in Breast Cancer, 20-54 (Kelly K Hunt et al., ed., 2001); Kochanek *et al.*, 49 *Nat'l. Vital Statistics Reports* 1, 14 (2001). Breast 30 cancer is extremely rare in women younger than 20 and is very rare in women under 30. The incidence of breast cancer rises with age and becomes significant by age 50. White Non-Hispanic women have the highest incidence rate for breast cancer and Korean women have the lowest. Increased prevalence of the genetic mutations BRCA1 and BRCA2 that

promote breast and other cancers are found in Ashkenazi Jews. African American women have the highest mortality rate for breast cancer among these same groups (31 per 100,000), while Chinese women have the lowest at 11 per 100,000. Although men can get breast cancer, this is extremely rare. In the United States it is estimated there will be 5 212,600 new cases of breast cancer and 40,200 deaths due to breast cancer in 2003. (American Cancer Society Website: cancer.org at the world wide web). With the exception of those cases with associated genetic factors, precise causes of breast cancer are not known.

In the treatment of breast cancer, there is considerable emphasis on detection and 10 risk assessment because early and accurate staging of breast cancer has a significant impact on survival. For example, breast cancer detected at an early stage (stage T0, discussed below) has a five-year survival rate of 92%. Conversely, if the cancer is not detected until a late stage (i.e., stage T4 (IV)), the five-year survival rate is reduced to 13%. AJCC Cancer Staging Handbook pp. 164-65 (Irvin D. Fleming *et al.* eds., 5th ed. 15 1998). Some detection techniques, such as mammography and biopsy, involve increased discomfort, expense, and/or radiation, and are prescribed only to patients with an increased risk of breast cancer.

Current methods for predicting or detecting breast cancer risk are not optimal. One method for predicting the relative risk of breast cancer is by examining a patient's risk 20 factors and pursuing aggressive diagnostic and treatment regimens for high risk patients. A patient's risk of breast cancer has been positively associated with increasing age, nulliparity, family history of breast cancer, personal history of breast cancer, early menarche, late menopause, late age of first full term pregnancy, prior proliferative breast disease, irradiation of the breast at an early age and a personal history of malignancy. 25 Lifestyle factors such as fat consumption, alcohol consumption, education, and socioeconomic status have also been associated with an increased incidence of breast cancer although a direct cause and effect relationship has not been established. While these risk factors are statistically significant, their weak association with breast cancer limits their usefulness. Most women who develop breast cancer have none of the risk 30 factors listed above, other than the risk that comes with growing older. NIH Publication No. 00-1556 (2000).

Current screening methods for detecting cancer, such as breast self exam, ultrasound, and mammography have drawbacks that reduce their effectiveness or prevent

their widespread adoption. Breast self exams, while useful, are unreliable for the detection of breast cancer in the initial stages where the tumor is small and difficult to detect by palpation. Ultrasound measurements require skilled operators at an increased expense. Mammography, while sensitive, is subject to over diagnosis in the detection of lesions that 5 have questionable malignant potential. There is also the fear of the radiation used in mammography because prior chest radiation is a factor associated with an increased incidence of breast cancer.

At this time, there are no adequate methods of breast cancer prevention. The current methods of breast cancer prevention involve prophylactic mastectomy 10 (mastectomy performed before cancer diagnosis) and chemoprevention (chemotherapy before cancer diagnosis) which are drastic measures that limit their adoption even among women with increased risk of breast cancer. Bevers, *supra*.

A number of genetic markers have been associated with breast cancer. Examples 15 of these markers include carcinoembryonic antigen (CEA) (Mughal *et al.*, *JAMA* 249:1881 (1983)), MUC-1 (Frische and Liu, *J. Clin. Ligand* 22:320 (2000)), HER-2/neu (Haris *et al.*, *Proc.Am.Soc.Clin.Oncology* 15:A96 (1996)), uPA, PAI-1, LPA, LPC, RAK and BRCA (Esteva and Fritsche, *Serum and Tissue Markers for Breast Cancer*, in Breast 20 Cancer, 286-308 (2001)). These markers have problems with limited sensitivity, low correlation, and false negatives which limit their use for initial diagnosis. For example, while the BRCA1 gene mutation is useful as an indicator of an increased risk for breast cancer, it has limited use in cancer diagnosis because only 6.2 % of breast cancers are BRCA1 positive. Malone *et al.*, *JAMA* 279:922 (1998). See also, Mewman *et al.*, *JAMA* 279:915 (1998) (correlation of only 3.3%).

There are four primary classifications of breast cancer varying by the site of origin 25 and the extent of disease development.

- I. Ductal carcinoma in situ (DCIS): Malignant transformation of ductal epithelial cells that remain in their normal position. DCIS is a purely localized disease, incapable of metastasis.
- II. Invasive ductal carcinoma (IDC): Malignancy of the ductal epithelial cells breaking through the basal membrane and into the supporting tissue of the breast. IDC may eventually spread elsewhere in the body.
- III. Lobular carcinoma in situ (LCIS): Malignancy arising in a single lobule of the breast that fail to extend through the lobule wall, it generally remains localized.

IV. Infiltrating lobular carcinoma (ILC): Malignancy arising in a single lobule of the breast and invading directly through the lobule wall into adjacent tissues.

By virtue of its invasion beyond the lobule wall, ILC may penetrate lymphatics and blood vessels and spread to distant sites.

5 For purpose of determining prognosis and treatment, these four breast cancer types have been staged according to the size of the primary tumor (T), the involvement of lymph nodes (N), and the presence of metastasis (M). Although DCIS by definition represents localized stage I disease, the other forms of breast cancer may range from stage II to stage IV. There are additional prognostic factors that further serve to guide surgical and medical
10 intervention. The most common ones are total number of lymph nodes involved, ER (estrogen receptor) status, Her2/neu receptor status and histologic grades.

Breast cancers are diagnosed into the appropriate stage categories recognizing that different treatments are more effective for different stages of cancer. Stage TX indicates that primary tumor cannot be assessed (i.e., tumor was removed or breast tissue was
15 removed). Stage T0 is characterized by abnormalities such as hyperplasia but with no evidence of primary tumor. Stage Tis is characterized by carcinoma in situ, intraductal carcinoma, lobular carcinoma in situ, or Paget's disease of the nipple with no tumor. Stage T1 (I) is characterized as having a tumor of 2 cm or less in the greatest dimension. Within stage T1, Tmic indicates microinvasion of 0.1 cm or less, T1a indicates a tumor of
20 between 0.1 to 0.5 cm, T1b indicates a tumor of between 0.5 to 1 cm, and T1c indicates tumors of between 1 cm to 2 cm. Stage T2 (II) is characterized by tumors from 2 cm to 5 cm in the greatest dimension. Tumors greater than 5 cm in size are classified as stage T3 (III). Stage T4 (IV) indicates a tumor of any size with extension to the chest wall or skin. Within stage T4, T4a indicates extension of the tumor to the chest wall, T4b indicates
25 edema or ulceration of the skin of the breast or satellite skin nodules confined to the same breast, T4c indicates a combination of T4a and T4b, and T4d indicates inflammatory carcinoma. AJCC Cancer Staging Handbook pp. 159-70 (Irvin D. Fleming *et al.* eds., 5th ed. 1998). In addition to standard staging, breast tumors may be classified according to their estrogen receptor and progesterone receptor protein status. Fisher *et al.*, *Breast
30 Cancer Research and Treatment* 7:147 (1986). Additional pathological status, such as HER2/neu status may also be useful. Thor *et al.*, *J.Nat'l.Cancer Inst.* 90:1346 (1998); Paik *et al.*, *J.Nat'l.Cancer Inst.* 90:1361 (1998); Hutchins *et al.*,

Proc.Am.Soc.Clin.Oncology 17:A2 (1998); and Simpson *et al.*, *J.Clin.Oncology* 18:2059 (2000).

In addition to the staging of the primary tumor, breast cancer metastases to regional lymph nodes may be staged. Stage NX indicates that the lymph nodes cannot be assessed (e.g., previously removed). Stage N0 indicates no regional lymph node metastasis. Stage N1 indicates metastasis to movable ipsilateral axillary lymph nodes. Stage N2 indicates metastasis to ipsilateral axillary lymph nodes fixed to one another or to other structures. Stage N3 indicates metastasis to ipsilateral internal mammary lymph nodes. *Id.*

Stage determination has potential prognostic value and provides criteria for designing optimal therapy. Simpson *et al.*, *J. Clin. Oncology* 18:2059 (2000). Generally, pathological staging of breast cancer is preferable to clinical staging because the former gives a more accurate prognosis. However, clinical staging would be preferred if it were as accurate as pathological staging because it does not depend on an invasive procedure to obtain tissue for pathological evaluation. Staging of breast cancer would be improved by detecting new markers in cells, tissues, or bodily fluids which could differentiate between different stages of invasion. Progress in this field will allow more rapid and reliable method for treating breast cancer patients.

Treatment of breast cancer is generally decided after an accurate staging of the primary tumor. Primary treatment options include breast conserving therapy (lumpectomy, breast irradiation, and surgical staging of the axilla), and modified radical mastectomy. Additional treatments include chemotherapy, regional irradiation, and, in extreme cases, terminating estrogen production by ovarian ablation.

Until recently, the customary treatment for all breast cancer was mastectomy. Fonseca *et al.*, *Annals of Internal Medicine* 127:1013 (1997). However, recent data indicate that less radical procedures may be equally effective, in terms of survival, for early stage breast cancer. Fisher *et al.*, *J. of Clinical Oncology* 16:441 (1998). The treatment options for a patient with early stage breast cancer (i.e., stage Tis) may be breast-sparing surgery followed by localized radiation therapy at the breast. Alternatively, mastectomy optionally coupled with radiation or breast reconstruction may be employed. These treatment methods are equally effective in the early stages of breast cancer.

Patients with stage I and stage II breast cancer require surgery with chemotherapy and/or hormonal therapy. Surgery is of limited use in stage III and stage IV patients.

Thus, these patients are better candidates for chemotherapy and radiation therapy with surgery limited to biopsy to permit initial staging or subsequent restaging because cancer is rarely curative at this stage of the disease. AJCC Cancer Staging Handbook 84, 164-65 (Irvin D. Fleming *et al.* eds., 5th ed.1998).

- 5 In an effort to provide more treatment options to patients, efforts are underway to define an earlier stage of breast cancer with low recurrence which could be treated with lumpectomy without postoperative radiation treatment. While a number of attempts have been made to classify early stage breast cancer, no consensus recommendation on postoperative radiation treatment has been obtained from these studies. Page *et al.*,
10 *Cancer* 75:1219 (1995); Fisher *et al.*, *Cancer* 75:1223 (1995); Silverstein *et al.*, *Cancer* 77:2267 (1996).

Cancer of the ovaries is the fourth most common cause of cancer death in women in the United States, with more than 23,000 new cases and roughly 14,000 deaths predicted for the year 2001. Shridhar, V. *et al.*, *Cancer Res.* 61(15):5895-904 (2001);
15 Memarzadeh, S. & Berek, J. S., *J. Reprod. Med.* 46(7):621-29 (2001). The incidence of ovarian cancer is of serious concern worldwide, with an estimated 191,000 new cases predicted annually. Runnebaum, I. B. & Stickeler, E., *J.Cancer Res. Clin. Oncol.* 127(2):73-79 (2001). These numbers continue to rise today. In the United States alone, it is estimated there will be 25,400 new cases of ovarian cancer, and 14,300 deaths
20 due to ovarian cancer in 2003. (American Cancer Society Website:
<http://www.cancer.org>). Unfortunately, women with ovarian cancer are typically asymptomatic until the disease has metastasized. Because effective screening for ovarian cancer is not available, roughly 70% of women diagnosed have an advanced stage of the cancer with a five-year survival rate of ~25-30%. Memarzadeh, S. & Berek, J. S., *supra*;
25 Nunns, D. *et al.*, *Obstet. Gynecol. Surv.* 55(12):746-51. Conversely, women diagnosed with early stage ovarian cancer enjoy considerably higher survival rates. Werness, B. A. & Eltabbakh, G. H., *Int'l. J. Gynecol. Pathol.* 20(1):48-63 (2001). Although our understanding of the etiology of ovarian cancer is incomplete, the results of extensive research in this area point to a combination of age, genetics, reproductive, and
30 dietary/environmental factors. Age is a key risk factor in the development of ovarian cancer: while the risk for developing ovarian cancer before the age of 30 is slim, the incidence of ovarian cancer rises linearly between ages 30 to 50, increasing at a slower rate thereafter, with the highest incidence being among septagenarian women. Jeanne M.

Schilder *et al.*, Hereditary Ovarian Cancer: Clinical Syndromes and Management, in Ovarian Cancer 182 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001).

- With respect to genetic factors, a family history of ovarian cancer is the most significant risk factor in the development of the disease, with that risk depending on the
- 5 number of affected family members, the degree of their relationship to the woman, and which particular first degree relatives are affected by the disease. *Id.* Mutations in several genes have been associated with ovarian cancer, including BRCA1 and BRCA2, both of which play a key role in the development of breast cancer, as well as hMSH2 and hMLH1, both of which are associated with hereditary non-polyposis colon cancer. Katherine Y.
- 10 Look, Epidemiology, Etiology, and Screening of Ovarian Cancer, in Ovarian Cancer 169, 171-73 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001). BRCA1, located on chromosome 17, and BRCA2, located on chromosome 13, are tumor suppressor genes implicated in DNA repair; mutations in these genes are linked to roughly 10% of ovarian cancers. *Id.* at 171-72; Schilder *et al.*, *supra* at 185-86. hMSH2 and hMLH1 are
- 15 associated with DNA mismatch repair, and are located on chromosomes 2 and 3, respectively; it has been reported that roughly 3% of hereditary ovarian carcinomas are due to mutations in these genes. Look, *supra* at 173; Schilder *et al.*, *supra* at 184, 188-89.

Reproductive factors have also been associated with an increased or reduced risk of ovarian cancer. Late menopause, nulliparity, and early age at menarche have all been

20 linked with an elevated risk of ovarian cancer. Schilder *et al.*, *supra* at 182. One theory hypothesizes that these factors increase the number of ovulatory cycles over the course of a woman's life, leading to "incessant ovulation," which is thought to be the primary cause of mutations to the ovarian epithelium. *Id.*; Laura J. Havrilesky & Andrew Berchuck, Molecular Alterations in Sporadic Ovarian Cancer, in Ovarian Cancer 25 (Stephen C.

25 Rubin & Gregory P. Sutton eds., 2d ed. 2001). The mutations may be explained by the fact that ovulation results in the destruction and repair of that epithelium, necessitating increased cell division, thereby increasing the possibility that an undetected mutation will occur. *Id.* Support for this theory may be found in the fact that pregnancy, lactation, and the use of oral contraceptives, all of which suppress ovulation, confer a protective effect

30 with respect to developing ovarian cancer. *Id.*

Among dietary/environmental factors, there would appear to be an association between high intake of animal fat or red meat and ovarian cancer, while the antioxidant Vitamin A, which prevents free radical formation and also assists in maintaining normal

cellular differentiation, may offer a protective effect. Look, *supra* at 169. Reports have also associated asbestos and hydrous magnesium trisilicate (talc), the latter of which may be present in diaphragms and sanitary napkins. *Id.* at 169-70.

Current screening procedures for ovarian cancer, while of some utility, are quite limited in their diagnostic ability, a problem that is particularly acute at early stages of cancer progression when the disease is typically asymptomatic yet is most readily treatable. Walter J. Burdette, Cancer: Etiology, Diagnosis, and Treatment 166 (1998); Memarzadeh & Berek, *supra*; Runnebaum & Stickeler, *supra*; Werness & Eltabakh, *supra*. Commonly used screening tests include biannual rectovaginal pelvic examination, radioimmunoassay to detect the CA-125 serum tumor marker, and transvaginal ultrasonography. Burdette, *supra* at 166.

Pelvic examination has failed to yield adequate numbers of early diagnoses, and the other methods are not sufficiently accurate. *Id.* One study reported that only 15% of patients who suffered from ovarian cancer were diagnosed with the disease at the time of their pelvic examination. Look, *supra* at 174. Moreover, the CA-125 test is prone to giving false positives in pre-menopausal women and has been reported to be of low predictive value in post-menopausal women. *Id.* at 174-75. Although transvaginal ultrasonography is now the preferred procedure for screening for ovarian cancer, it is unable to distinguish reliably between benign and malignant tumors, and also cannot locate primary peritoneal malignancies or ovarian cancer if the ovary size is normal. Schilder *et al.*, *supra* at 194-95. While genetic testing for mutations of the BRCA1, BRCA2, hMSH2, and hMLH1 genes is now available, these tests may be too costly for some patients and may also yield false negative or indeterminate results. Schilder *et al.*, *supra* at 191-94.

The staging of ovarian cancer, which is accomplished through surgical exploration, is crucial in determining the course of treatment and management of the disease. AJCC Cancer Staging Handbook 187 (Irvin D. Fleming *et al.* eds., 5th ed. 1998); Burdette, *supra* at 170; Memarzadeh & Berek, *supra*; Shridhar *et al.*, *supra*. Staging is performed by reference to the classification system developed by the International Federation of Gynecology and Obstetrics. David H. Moore, Primary Surgical Management of Early Epithelial Ovarian Carcinoma, in Ovarian Cancer 203 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001); Fleming *et al.* eds., *supra* at 188. Stage I ovarian cancer is characterized by tumor growth that is limited to the ovaries and is comprised of three

substages. *Id.* In substage IA, tumor growth is limited to one ovary, there is no tumor on the external surface of the ovary, the ovarian capsule is intact, and no malignant cells are present in ascites or peritoneal washings. *Id.* Substage IB is identical to A1, except that tumor growth is limited to both ovaries. *Id.* Substage IC refers to the presence of tumor growth limited to one or both ovaries, and also includes one or more of the following characteristics: capsule rupture, tumor growth on the surface of one or both ovaries, and malignant cells present in ascites or peritoneal washings. *Id.*

Stage II ovarian cancer refers to tumor growth involving one or both ovaries, along with pelvic extension. *Id.* Substage IIA involves extension and/or implants on the uterus and/or fallopian tubes, with no malignant cells in the ascites or peritoneal washings, while substage IIB involves extension into other pelvic organs and tissues, again with no malignant cells in the ascites or peritoneal washings. *Id.* Substage IIC involves pelvic extension as in IIA or IIB, but with malignant cells in the ascites or peritoneal washings. *Id.*

Stage III ovarian cancer involves tumor growth in one or both ovaries, with peritoneal metastasis beyond the pelvis confirmed by microscope and/or metastasis in the regional lymph nodes. *Id.* Substage IIIA is characterized by microscopic peritoneal metastasis outside the pelvis, with substage IIIB involving macroscopic peritoneal metastasis outside the pelvis 2 cm or less in greatest dimension. *Id.* Substage IIIC is identical to IIIB, except that the metastasis is greater than 2 cm in greatest dimension and may include regional lymph node metastasis. *Id.* Lastly, Stage IV refers to the presence of distant metastasis, excluding peritoneal metastasis. *Id.*

While surgical staging is currently the benchmark for assessing the management and treatment of ovarian cancer, it suffers from considerable drawbacks, including the invasiveness of the procedure, the potential for complications, as well as the potential for inaccuracy. Moore, *supra* at 206-208, 213. In view of these limitations, attention has turned to developing alternative staging methodologies through understanding differential gene expression in various stages of ovarian cancer and by obtaining various biomarkers to help better assess the progression of the disease. Vartiainen, J. *et al.*, *Int'l J. Cancer*, 95(5):313-16 (2001); Shridhar *et al.* *supra*; Baekelandt, M. *et al.*, *J. Clin. Oncol.* 18(22):3775-81.

The treatment of ovarian cancer typically involves a multiprong attack, with surgical intervention serving as the foundation of treatment. Dennis S. Chi & William J.

Hoskins, *Primary Surgical Management of Advanced Epithelial Ovarian Cancer*, in Ovarian Cancer 241 (Stephen C. Rubin & Gregory P. Sutton eds., 2d ed. 2001). For example, in the case of epithelial ovarian cancer, which accounts for ~90% of cases of ovarian cancer, treatment typically consists of: (1) cytoreductive surgery, including total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, and lymphadenectomy, followed by (2) adjuvant chemotherapy with paclitaxel and either cisplatin or carboplatin. Eltabbakh, G.H. & Awtrey, C.S., *Expert Op. Pharmacother.* 2(10):109-24. Despite a clinical response rate of 80% to the adjuvant therapy, most patients experience tumor recurrence within three years of treatment. *Id.* Certain patients may undergo a second cytoreductive surgery and/or second-line chemotherapy.

Memarzadeh & Berek, *supra*.

From the foregoing, it is clear that procedures used for detecting, diagnosing, monitoring, staging, prognosticating, and preventing the recurrence of ovarian cancer are of critical importance to the outcome of the patient. Moreover, current procedures, while helpful in each of these analyses, are limited by their specificity, sensitivity, invasiveness, and/or their cost. As such, highly specific and sensitive procedures that would operate by way of detecting novel markers in cells, tissues, or bodily fluids, with minimal invasiveness and at a reasonable cost, would be highly desirable.

As discussed above, each of the methods for diagnosing and staging ovarian, pancreatic or breast cancer is limited by the technology employed. Accordingly, there is need for sensitive molecular and cellular markers for the detection of ovarian, pancreatic or breast cancer. There is a need for molecular markers for the accurate staging, including clinical and pathological staging, of ovarian, pancreatic or breast cancers to optimize treatment methods. Finally, there is a need for sensitive molecular and cellular markers to monitor the progress of cancer treatments, including markers that can detect recurrence of ovarian, pancreatic or breast cancers following remission.

The present invention provides alternative methods of treating ovarian, pancreatic or breast cancer that overcome the limitations of conventional therapeutic methods as well as offer additional advantages that will be apparent from the detailed description below.

Growth and metastasis of solid tumors are also dependent on angiogenesis.

Folkman, J., 1986, *Cancer Research*, 46, 467-473; Folkman, J., 1989, *Journal of the National Cancer Institute*, 82, 4-6. It has been shown, for example, that tumors which enlarge to greater than 2 mm must obtain their own blood supply and do so by inducing

the growth of new capillary blood vessels. Once these new blood vessels become embedded in the tumor, they provide a means for tumor cells to enter the circulation and metastasize to distant sites such as liver, lung or bone. Weidner, N., *et al.*, 1991, *The New England Journal of Medicine*, 324(1), 1-8.

5 Angiogenesis, defined as the growth or sprouting of new blood vessels from existing vessels, is a complex process that primarily occurs during embryonic development. The process is distinct from vasculogenesis, in that the new endothelial cells lining the vessel arise from proliferation of existing cells, rather than differentiating from stem cells. The process is invasive and dependent upon proteolysis of the extracellular
10 matrix (ECM), migration of new endothelial cells, and synthesis of new matrix components. Angiogenesis occurs during embryogenic development of the circulatory system; however, in adult humans, angiogenesis only occurs as a response to a pathological condition (except during the reproductive cycle in women).

Under normal physiological conditions in adults, angiogenesis takes place only in
15 very restricted situations such as hair growth and wounding healing. Auerbach, W. and Auerbach, R., 1994, *Pharmacol Ther.* 63(3):265-311; Ribatti et al., 1991, *Haematologica* 76(4):311-20; Risau, 1997, *Nature* 386(6626):671-4. Angiogenesis progresses by a stimulus which results in the formation of a migrating column of endothelial cells. Proteolytic activity is focused at the advancing tip of this "vascular sprout", which breaks
20 down the ECM sufficiently to permit the column of cells to infiltrate and migrate. Behind the advancing front, the endothelial cells differentiate and begin to adhere to each other, thus forming a new basement membrane. The cells then cease proliferation and finally define a lumen for the new arteriole or capillary.

Unregulated angiogenesis has gradually been recognized to be responsible for a
25 wide range of disorders, including, but not limited to, cancer, cardiovascular disease, rheumatoid arthritis, psoriasis and diabetic retinopathy. Folkman, 1995, *Nat Med* 1(1):27-31; Isner, 1999, *Circulation* 99(13): 1653-5; Koch, 1998, *Arthritis Rheum* 41(6):951-62; Walsh, 1999, *Rheumatology* (Oxford) 38(2):103-12; Ware and Simons, 1997, *Nat Med* 3(2): 158-64.

30 Of particular interest is the observation that angiogenesis is required by solid tumors for their growth and metastases. Folkman, 1986 *supra*; Folkman 1990, *J Natl. Cancer Inst.*, 82(1) 4-6; Folkman, 1992, *Semin Cancer Biol* 3(2):65-71; Zetter, 1998, *Annu Rev Med* 49:407-24. A tumor usually begins as a single aberrant cell which can proliferate

only to a size of a few cubic millimeters due to the distance from available capillary beds, and it can stay 'dormant' without further growth and dissemination for a long period of time. Some tumor cells then switch to the angiogenic phenotype to activate endothelial cells, which proliferate and mature into new capillary blood vessels. These newly formed
5 blood vessels not only allow for continued growth of the primary tumor, but also for the dissemination and recolonization of metastatic tumor cells. The precise mechanisms that control the angiogenic switch is not well understood, but it is believed that neovascularization of tumor mass results from the net balance of a multitude of angiogenesis stimulators and inhibitors Folkman, 1995, *supra*.

10 One of the most potent angiogenesis inhibitors is endostatin identified by O'Reilly and Folkman. O'Reilly et al., 1997, *Cell* 88(2):277-85; O'Reilly et al., 1994, *Cell* 79(2):3 15-28. Its discovery was based on the phenomenon that certain primary tumors can inhibit the growth of distant metastases. O'Reilly and Folkman hypothesized that a primary tumor initiates angiogenesis by generating angiogenic stimulators in excess of inhibitors.
15 However, angiogenic inhibitors, by virtue of their longer half life in the circulation, reach the site of a secondary tumor in excess of the stimulators. The net result is the growth of primary tumor and inhibition of secondary tumor. Endostatin is one of a growing list of such angiogenesis inhibitors produced by primary tumors. It is a proteolytic fragment of a larger protein: endostatin is a 20 kDa fragment of collagen XVIII (amino acid H1132-K1315 in murine collagen XVIII). Endostatin has been shown to specifically inhibit
20 endothelial cell proliferation in vitro and block angiogenesis in vivo. More importantly, administration of endostatin to tumor-bearing mice leads to significant tumor regression, and no toxicity or drug resistance has been observed even after multiple treatment cycles. Boehm et al., 1997, *Nature* 390(6658):404-407. The fact that endostatin targets genetically
25 stable endothelial cells and inhibits a variety of solid tumors makes it a very attractive candidate for anticancer therapy. Fidler and Ellis, 1994, *Cell* 79(2):185-8; Gastl et al., 1997, *Oncology* 54(3):177-84; Hinsbergh et al., 1999, *Ann Oncol* 10 Suppl 4:60-3. In addition, angiogenesis inhibitors have been shown to be more effective when combined with radiation and chemotherapeutic agents. Klement, 2000, *J. Clin Invest*, 105(8) R15-
30 24. Browder, 2000, *Cancer Res.* 6-(7) 1878-86, Arap et al., 1998, *Science* 279(5349):377-80; Mauceri et al., 1998, *Nature* 394(6690):287-91.

SUMMARY OF THE INVENTION

The present invention solves many needs in the art by providing nucleic acid molecules, polypeptides and antibodies thereto, variants and derivatives of the nucleic acids and polypeptides, and agonists and antagonists thereto that may be used to identify,

- 5 diagnose, monitor, stage, image and treat breast cancer and/or non-cancerous disease states in breast; identify and monitor breast tissue; and identify and design agonists and antagonists of polypeptides of the invention. The invention also provides gene therapy, methods for producing transgenic animals and cells, and methods for producing engineered breast tissue for treatment and research.

10 One aspect of the present invention relates to nucleic acid molecules that are specific to breast cells, breast tissue and/or the breast organ. These breast specific nucleic acids (BSNAs) may be a naturally occurring cDNA, genomic DNA, RNA, or a fragment of one of these nucleic acids, or may be a non-naturally occurring nucleic acid molecule. If the BSNA is genomic DNA, then the BSNA is a breast specific gene (BSG). If the
15 BSNA is RNA, then it is a breast specific transcript encoded by a BSG. Due to alternative splicing and transcriptional modification one BSG may encode for multiple breast specific RNAs. In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that is specific to breast. More preferred is a nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 96-232. In another preferred
20 embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-95. For the BSNA sequences listed herein, DEX0452_001.nt.1 corresponds to SEQ ID NO: 1. For sequences with multiple splice variants, the parent sequence DEX0452_001.nt.1, will be followed by DEX0452_001.nt.2, etc. for each splice variant. The sequences off the corresponding peptides are listed as DEX0452_001.aa.1, etc. For
25 the mapping of all of the nucleotides and peptides, see the table in the Example 1 section below.

This aspect of the present invention also relates to nucleic acid molecules that selectively hybridize or exhibit substantial sequence similarity to nucleic acid molecules encoding a Breast Specific Protein (BSP), or that selectively hybridize or exhibit

- 30 substantial sequence similarity to a BSNA. In one embodiment of the present invention the nucleic acid molecule comprises an allelic variant of a nucleic acid molecule encoding a BSP, or an allelic variant of a BSNA. In another embodiment, the nucleic acid molecule

comprises a part of a nucleic acid sequence that encodes a BSP or a part of a nucleic acid sequence of a BSNA.

In addition, this aspect of the present invention relates to a nucleic acid molecule further comprising one or more expression control sequences controlling the transcription 5 and/or translation of all or a part of a BSNA or the transcription and/or translation of a nucleic acid molecule that encodes all or a fragment of a BSP.

Another aspect of the present invention relates to vectors and/or host cells comprising a nucleic acid molecule of this invention. In a preferred embodiment, the nucleic acid molecule of the vector and/or host cell encodes all or a fragment of a BSP. In 10 another preferred embodiment, the nucleic acid molecule of the vector and/or host cell comprises all or a part of a BSNA. Vectors and host cells of the present invention are useful in the recombinant production of polypeptides, particularly BSPs of the present invention.

Another aspect of the present invention relates to polypeptides encoded by a 15 nucleic acid molecule of this invention. The polypeptide may comprise either a fragment or a full-length protein. In a preferred embodiment, the polypeptide is a BSP. However, this aspect of the present invention also relates to mutant proteins (muteins) of BSPs, fusion proteins of which a portion is a BSP, and proteins and polypeptides encoded by allelic variants of a BSNA as provided herein.

A further aspect of the present invention is a novel splice variant which encodes an amino acid sequence that provides a novel region to be targeted for the generation of reagents that can be used in the detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or function. This information can be used to directly or indirectly 25 facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

Another aspect of the present invention relates to antibodies and other binders that specifically bind to a polypeptide of the instant invention. Accordingly antibodies or 30 binders of the present invention specifically bind to BSPs, muteins, fusion proteins, and/or homologous proteins or polypeptides encoded by allelic variants of a BSNA as provided herein.

Another aspect of the present invention relates to agonists and antagonists of the nucleic acid molecules and polypeptides of this invention. The agonists and antagonists of the instant invention may be used to treat breast cancer and non-cancerous disease states in breast and to produce engineered breast tissue.

5 Another aspect of the present invention relates to methods for using the nucleic acid molecules to detect or amplify nucleic acid molecules that have similar or identical nucleic acid sequences compared to the nucleic acid molecules described herein. Such methods are useful in identifying, diagnosing, monitoring, staging, imaging and treating breast cancer and/or non-cancerous disease states in breast. Such methods are also useful
10 10 in identifying and/or monitoring breast tissue. In addition, measurement of levels of one or more of the nucleic acid molecules of this invention may be useful as a diagnostic as part of a panel in combination with known other markers, particularly those described in the breast cancer background section above.

Another aspect of the present invention relates to use of the nucleic acid molecules
15 15 of this invention in gene therapy, for producing transgenic animals and cells, and for producing engineered breast tissue for treatment and research.

Another aspect of the present invention relates to methods for detecting polypeptides of this invention, preferably using antibodies thereto. Such methods are useful to identify, diagnose, monitor, stage, image and treat breast cancer and non-
20 20 cancerous disease states in breast. In addition, measurement of levels of one or more of the polypeptides of this invention may be useful to identify, diagnose, monitor, stage, and/or image breast cancer in combination with known other markers, particularly those described in the breast cancer background section above. The polypeptides of the present invention can also be used to identify and/or monitor breast tissue, and to produce
25 25 engineered breast tissue.

Yet another aspect of the present invention relates to a computer readable means of storing the nucleic acid and amino acid sequences of the invention. The records of the computer readable means can be accessed for reading and displaying of sequences for comparison, alignment and ordering of the sequences of the invention to other sequences.
30 30 In addition, the computer records regarding the nucleic acid and/or amino acid sequences and/or measurements of their levels may be used alone or in combination with other markers to diagnose breast related diseases.

DETAILED DESCRIPTION OF THE INVENTION

Definitions and General Techniques

Unless otherwise defined herein, scientific and technical terms used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. Generally, nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well known and commonly used in the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. See, e.g., Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Press (2001); Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2000); Ausubel *et al.*, Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology – 4th Ed., Wiley & Sons (1999); Harlow and Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1990); and Harlow and Lane, Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1999).

Enzymatic reactions and purification techniques are performed according to manufacturer's specifications, as commonly accomplished in the art or as described herein. The nomenclatures used in connection with, and the laboratory procedures and techniques of, analytical chemistry, synthetic organic chemistry, and medicinal and pharmaceutical chemistry described herein are those well known and commonly used in the art. Standard techniques are used for chemical syntheses, chemical analyses, pharmaceutical preparation, formulation, and delivery, and treatment of patients.

The following terms, unless otherwise indicated, shall be understood to have the following meanings:

A "nucleic acid molecule" of this invention refers to a polymeric form of nucleotides and includes both sense and antisense strands of RNA, cDNA, genomic DNA,

and synthetic forms and mixed polymers of the above. A nucleotide refers to a ribonucleotide, deoxynucleotide or a modified form of either type of nucleotide. A "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." The term "nucleic acid molecule" usually refers to a molecule of at least 10 bases in length, unless otherwise specified. The term includes single- and double-stranded forms of DNA. In addition, a polynucleotide may include either or both naturally occurring and modified nucleotides linked together by naturally occurring and/or non-naturally occurring nucleotide linkages.

Nucleotides are represented by single letter symbols in nucleic acid molecule sequences. The following table lists symbols identifying nucleotides or groups of nucleotides which may occupy the symbol position on a nucleic acid molecule. See Nomenclature Committee of the International Union of Biochemistry (NC-IUB), Nomenclature for incompletely specified bases in nucleic acid sequences, Recommendations 1984., *Eur J Biochem.* 150(1):1-5 (1985).

Symbol	Meaning	Group/Origin of Designation	Complementary Symbol
a	a	Adenine	t/u
g	g	Guanine	c
c	c	Cytosine	g
t	t	Thymine	a
u	u	Uracil	a
r	g or a	puRine	y
y	t/u or c	pYrimidine	r
m	a or c	aMino	k
k	g or t/u	Keto	m
s	g or c	Strong interactions 3H-bonds	w
w	a or t/u	Weak interactions 2H-bonds	s
b	g or c or t/u	not a	v
d	a or g or t/u	not c	h
h	a or c or t/u	not g	d
v	a or g or c	not t, not u	b
n	a or g or c or t/u, unknown, or other	aNy	n

The nucleic acid molecules may be modified chemically or biochemically or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example, labels, methylation, substitution of one or more of the naturally occurring nucleotides with an analog, 20 internucleotide modifications such as uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoramidates, carbamates, etc.), charged linkages (e.g.,

phosphorothioates, phosphorodithioates, etc.), pendent moieties (*e.g.*, polypeptides), intercalators (*e.g.*, acridine, psoralen, etc.), chelators, alkylators, and modified linkages (*e.g.*, alpha anomeric nucleic acids, etc.) The term “nucleic acid molecule” also includes any topological conformation, including single-stranded, double-stranded, partially

- 5 duplexed, triplexed, hairpinned, circular and padlocked conformations. Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence via hydrogen bonding and other chemical interactions. Such molecules are known in the art and include, for example, those in which peptide linkages substitute for phosphate linkages in the backbone of the molecule.

10 A “gene” is defined as a nucleic acid molecule that comprises a nucleic acid sequence that encodes a polypeptide and the expression control sequences that surround the nucleic acid sequence that encodes the polypeptide. For instance, a gene may comprise a promoter, one or more enhancers, a nucleic acid sequence that encodes a polypeptide, downstream regulatory sequences and, possibly, other nucleic acid sequences
15 involved in regulation of the expression of an RNA. As is well known in the art, eukaryotic genes usually contain both exons and introns. The term “exon” refers to a nucleic acid sequence found in genomic DNA that is bioinformatically predicted and/or experimentally confirmed to contribute contiguous sequence to a mature mRNA transcript. The term “intron” refers to a nucleic acid sequence found in genomic DNA that
20 is predicted and/or confirmed to not contribute to a mature mRNA transcript, but rather to be “spliced out” during processing of the transcript.

A nucleic acid molecule or polypeptide is “derived” from a particular species if the nucleic acid molecule or polypeptide has been isolated from the particular species, or if the nucleic acid molecule or polypeptide is homologous to a nucleic acid molecule or polypeptide isolated from a particular species.
25

An “isolated” or “substantially pure” nucleic acid or polynucleotide (*e.g.*, an RNA, DNA or a mixed polymer) is one which is substantially separated from other cellular components that naturally accompany the native polynucleotide in its natural host cell, *e.g.*, ribosomes, polymerases, or genomic sequences with which it is naturally associated.
30 The term embraces a nucleic acid or polynucleotide that (1) has been removed from its naturally occurring environment, (2) is not associated with all or a portion of a polynucleotide in which the “isolated polynucleotide” is found in nature, (3) is operatively linked to a polynucleotide which it is not linked to in nature, (4) does not occur in nature

as part of a larger sequence or (5) includes nucleotides or internucleoside bonds that are not found in nature. The term “isolated” or “substantially pure” also can be used in reference to recombinant or cloned DNA isolates, chemically synthesized polynucleotide analogs, or polynucleotide analogs that are biologically synthesized by heterologous systems. The term “isolated nucleic acid molecule” includes nucleic acid molecules that are integrated into a host cell chromosome at a heterologous site, recombinant fusions of a native fragment to a heterologous sequence, recombinant vectors present as episomes or as integrated into a host cell chromosome.

A “part” of a nucleic acid molecule refers to a nucleic acid molecule that comprises a partial contiguous sequence of at least 10 bases of the reference nucleic acid molecule. Preferably, a part comprises at least 15 to 20 bases of a reference nucleic acid molecule. In theory, a nucleic acid sequence of 17 nucleotides is of sufficient length to occur at random less frequently than once in the three gigabase human genome, and thus provides a nucleic acid probe that can uniquely identify the reference sequence in a nucleic acid mixture of genomic complexity. A preferred part is one that comprises a nucleic acid sequence that can encode at least 6 contiguous amino acid sequences (fragments of at least 18 nucleotides) because they are useful in directing the expression or synthesis of peptides that are useful in mapping the epitopes of the polypeptide encoded by the reference nucleic acid. *See, e.g., Geysen et al., Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1984); and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. A part may also comprise at least 25, 30, 35 or 40 nucleotides of a reference nucleic acid molecule, or at least 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides of a reference nucleic acid molecule. A part of a nucleic acid molecule may comprise no other nucleic acid sequences. Alternatively, a part of a nucleic acid may comprise other nucleic acid sequences from other nucleic acid molecules.

The term “oligonucleotide” refers to a nucleic acid molecule generally comprising a length of 200 bases or fewer. The term often refers to single-stranded deoxyribonucleotides, but it can refer as well to single- or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others. Preferably, oligonucleotides are 10 to 60 bases in length and most preferably 12, 13, 14, 15, 16, 17, 18, 19 or 20 bases in length. Other preferred oligonucleotides are 25, 30, 35, 40, 45, 50, 55 or 60 bases in length. Oligonucleotides may be single-stranded, e.g. for use as probes

or primers, or may be double-stranded, e.g. for use in the construction of a mutant gene. Oligonucleotides of the invention can be either sense or antisense oligonucleotides. An oligonucleotide can be derivatized or modified as discussed above for nucleic acid molecules.

5 Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms. Initially, chemically synthesized DNAs
10 typically are obtained without a 5' phosphate. The 5' ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP. The 3' end of a chemically synthesized
15 oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5' phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5' phosphates of the other polynucleotide(s) prior to ligation.

20 The term "naturally occurring nucleotide" referred to herein includes naturally occurring deoxyribonucleotides and ribonucleotides. The term "modified nucleotides" referred to herein includes nucleotides with modified or substituted sugar groups and the like. The term "nucleotide linkages" referred to herein includes nucleotide linkages such as phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate,
25 phosphoroanilothioate, phosphoraniladate, phosphoroamidate, and the like. See e.g., LaPlanche *et al.* *Nucl. Acids Res.* 14:9081-9093 (1986); Stein *et al.* *Nucl. Acids Res.* 16:3209-3221 (1988); Zon *et al.* *Anti-Cancer Drug Design* 6:539-568 (1991); Zon *et al.*, in Eckstein (ed.) Oligonucleotides and Analogues: A Practical Approach, pp. 87-108, Oxford University Press (1991); Uhlmann and Peyman *Chemical Reviews* 90:543 (1990),
30 and U.S. Patent No. 5,151,510, the disclosure of which is hereby incorporated by reference in its entirety.

Unless specified otherwise, the left hand end of a polynucleotide sequence in sense orientation is the 5' end and the right hand end of the sequence is the 3' end. In addition,

the left hand direction of a polynucleotide sequence in sense orientation is referred to as the 5' direction, while the right hand direction of the polynucleotide sequence is referred to as the 3' direction. Further, unless otherwise indicated, each nucleotide sequence is set forth herein as a sequence of deoxyribonucleotides. It is intended, however, that the given
5 sequence be interpreted as would be appropriate to the polynucleotide composition: for example, if the isolated nucleic acid is composed of RNA, the given sequence intends ribonucleotides, with uridine substituted for thymidine.

The term "allelic variant" refers to one of two or more alternative naturally occurring forms of a gene, wherein each gene possesses a unique nucleotide sequence. In
10 a preferred embodiment, different alleles of a given gene have similar or identical biological properties.

The term "percent sequence identity" in the context of nucleic acid sequences refers to the residues in two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over a stretch of at
15 least about nine nucleotides, usually at least about 20 nucleotides, more usually at least about 24 nucleotides, typically at least about 28 nucleotides, more typically at least about 32 nucleotides, and preferably at least about 36 or more nucleotides. There are a number of different algorithms known in the art which can be used to measure nucleotide sequence identity. For instance, polynucleotide sequences can be compared using FASTA, Gap or
20 Bestfit, which are programs in Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin. FASTA, which includes, e.g., the programs FASTA2 and FASTA3, provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences (Pearson, *Methods Enzymol.* 183: 63-98 (1990); Pearson, *Methods Mol. Biol.* 132: 185-219 (2000); Pearson, *Methods Enzymol.* 25 266: 227-258 (1996); Pearson, *J. Mol. Biol.* 276: 71-84 (1998)). Unless otherwise specified, default parameters for a particular program or algorithm are used. For instance, percent sequence identity between nucleic acid sequences can be determined using FASTA with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) or using Gap with its default parameters as provided in GCG Version 6.1.
25 A reference to a nucleic acid sequence encompasses its complement unless otherwise specified. Thus, a reference to a nucleic acid molecule having a particular sequence should be understood to encompass its complementary strand, with its complementary sequence. The complementary strand is also useful, e.g., for antisense
30

therapy, double-stranded RNA (dsRNA) inhibition (RNAi), combination of triplex and antisense, hybridization probes and PCR primers.

In the molecular biology art, researchers use the terms "percent sequence identity", "percent sequence similarity" and "percent sequence homology" interchangeably. In this
5 application, these terms shall have the same meaning with respect to nucleic acid sequences only.

The term "substantial similarity" or "substantial sequence similarity," when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its
10 complementary strand), there is nucleotide sequence identity in at least about 50%, more preferably 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, and more preferably at least about 95-98% of the nucleotide bases, as measured by any well known algorithm of sequence identity, such as FASTA, BLAST or Gap, as discussed above.

15 Alternatively, substantial similarity exists between a first and second nucleic acid sequence when the first nucleic acid sequence or fragment thereof hybridizes to an antisense strand of the second nucleic acid, under selective hybridization conditions. Typically, selective hybridization will occur between the first nucleic acid sequence and an antisense strand of the second nucleic acid sequence when there is at least about 55%
20 sequence identity between the first and second nucleic acid sequences—preferably at least about 65%, more preferably at least about 75%, and most preferably at least about 90%—over a stretch of at least about 14 nucleotides, more preferably at least 17 nucleotides, even more preferably at least 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 or 100 nucleotides.

Nucleic acid hybridization will be affected by such conditions as salt
25 concentration, temperature, solvents, the base composition of the hybridizing species, length of the complementary regions, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. "Stringent hybridization conditions" and "stringent wash conditions" in the context of nucleic acid hybridization experiments depend upon a number of different physical
30 parameters. The most important parameters include temperature of hybridization, base composition of the nucleic acids, salt concentration and length of the nucleic acid. One having ordinary skill in the art knows how to vary these parameters to achieve a particular stringency of hybridization. In general, "stringent hybridization" is performed at about

25°C below the thermal melting point (T_m) for the specific DNA hybrid under a particular set of conditions. "Stringent washing" is performed at temperatures about 5°C lower than the T_m for the specific DNA hybrid under a particular set of conditions. The T_m is the temperature at which 50% of the target sequence hybridizes to a perfectly matched probe.

5 See Sambrook (1989), *supra*, p. 9.51.

The T_m for a particular DNA-DNA hybrid can be estimated by the formula:

$$T_m = 81.5^\circ\text{C} + 16.6 (\log_{10}[\text{Na}^+]) + 0.41 (\text{fraction G} + \text{C}) - 0.63 (\%) \text{ formamide} - (600/l) \text{ where } l \text{ is the length of the hybrid in base pairs.}$$

The T_m for a particular RNA-RNA hybrid can be estimated by the formula:

10 $T_m = 79.8^\circ\text{C} + 18.5 (\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) + 11.8 (\text{fraction G} + \text{C})^2 - 0.35 (\%) \text{ formamide} - (820/l).$

The T_m for a particular RNA-DNA hybrid can be estimated by the formula:

$$T_m = 79.8^\circ\text{C} + 18.5(\log_{10}[\text{Na}^+]) + 0.58 (\text{fraction G} + \text{C}) + 11.8 (\text{fraction G} + \text{C})^2 - 0.50 (\%) \text{ formamide} - (820/l).$$

15 In general, the T_m decreases by 1-1.5°C for each 1% of mismatch between two nucleic acid sequences. Thus, one having ordinary skill in the art can alter hybridization and/or washing conditions to obtain sequences that have higher or lower degrees of sequence identity to the target nucleic acid. For instance, to obtain hybridizing nucleic acids that contain up to 10% mismatch from the target nucleic acid sequence, 10-15°C would be subtracted from the calculated T_m of a perfectly matched hybrid, and then the hybridization and washing temperatures adjusted accordingly. Probe sequences may also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

20

25 An example of stringent hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100 complementary residues on a filter in a Southern or Northern blot or for screening a library is 50% formamide/6X SSC at 42°C for at least ten hours and preferably overnight (approximately 16 hours). Another example of stringent hybridization conditions is 6X SSC at 68°C without formamide for at least ten hours and preferably overnight. An example of moderate stringency hybridization conditions is 6X SSC at 55°C without formamide for at least ten hours and preferably overnight. An example of low stringency hybridization conditions for hybridization of complementary nucleic acid sequences having more than 100

30

complementary residues on a filter in a Southern or northern blot or for screening a library is 6X SSC at 42°C for at least ten hours. Hybridization conditions to identify nucleic acid sequences that are similar but not identical can be identified by experimentally changing the hybridization temperature from 68°C to 42°C while keeping the salt concentration constant (6X SSC), or keeping the hybridization temperature and salt concentration constant (e.g. 42°C and 6X SSC) and varying the formamide concentration from 50% to 0%. Hybridization buffers may also include blocking agents to lower background. These agents are well known in the art. See Sambrook *et al.* (1989), *supra*, pages 8.46 and 9.46-9.58. See also Ausubel (1992), *supra*, Ausubel (1999), *supra*, and Sambrook (2001), *supra*.

Wash conditions also can be altered to change stringency conditions. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (see Sambrook (1989), *supra*, for SSC buffer). Often the high stringency wash is preceded by a low stringency wash to remove excess probe. An exemplary medium stringency wash for duplex DNA of more than 100 base pairs is 1x SSC at 45°C for 15 minutes. An exemplary low stringency wash for such a duplex is 4x SSC at 40°C for 15 minutes. In general, signal-to-noise ratio of 2x or higher than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

As defined herein, nucleic acids that do not hybridize to each other under stringent conditions are still substantially similar to one another if they encode polypeptides that are substantially identical to each other. This occurs, for example, when a nucleic acid is created synthetically or recombinantly using a high codon degeneracy as permitted by the redundancy of the genetic code.

Hybridization conditions for nucleic acid molecules that are shorter than 100 nucleotides in length (e.g., for oligonucleotide probes) may be calculated by the formula:

$T_m = 81.5^\circ\text{C} + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\text{fraction G+C}) - (600/N)$, wherein N is change length and the $[\text{Na}^+]$ is 1 M or less. See Sambrook (1989), *supra*, p. 11.46. For hybridization of probes shorter than 100 nucleotides, hybridization is usually performed under stringent conditions (5-10°C below the T_m) using high concentrations (0.1-1.0 pmol/ml) of probe. *Id.* at p. 11.45. Determination of hybridization using mismatched probes, pools of degenerate probes or "guessmers," as well as hybridization solutions and methods for empirically determining hybridization conditions are well known in the art. See, e.g., Ausubel (1999), *supra*; Sambrook (1989), *supra*, pp. 11.45-11.57.

The term "digestion" or "digestion of DNA" refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan. For analytical purposes, typically, 1 µg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 µl of reaction buffer. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 µg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes. Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and are specified by commercial suppliers. Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

The term "ligation" refers to the process of forming phosphodiester bonds between two or more polynucleotides, which most often are double-stranded DNAs. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, e.g., Sambrook (1989), *supra*.

Genome-derived "single exon probes," are probes that comprise at least part of an exon ("reference exon") and can hybridize detectably under high stringency conditions to transcript-derived nucleic acids that include the reference exon but do not hybridize detectably under high stringency conditions to nucleic acids that lack the reference exon. Single exon probes typically further comprise, contiguous to a first end of the exon portion, a first intronic and/or intergenic sequence that is identically contiguous to the exon in the genome, and may contain a second intronic and/or intergenic sequence that is identically contiguous to the exon in the genome. The minimum length of genome-derived single exon probes is defined by the requirement that the exonic portion be of sufficient length to hybridize under high stringency conditions to transcript-derived nucleic acids, as discussed above. The maximum length of genome-derived single exon probes is defined by the requirement that the probes contain portions of no more than one exon. The single exon probes may contain priming sequences not found in contiguity with the rest of the probe sequence in the genome, which priming sequences are useful for PCR

and other amplification-based technologies. In another aspect, the invention is directed to single exon probes based on the BSNAs disclosed herein.

In one embodiment, the term "microarray" refers to a "nucleic acid microarray" having a substrate-bound plurality of nucleic acids, hybridization to each of the plurality of bound nucleic acids being separately detectable. The substrate can be solid or porous, planar or non-planar, unitary or distributed. Nucleic acid microarrays include all the devices so called in Schena (ed.), DNA Microarrays: A Practical Approach (Practical Approach Series), Oxford University Press (1999); *Nature Genet.* 21(1)(suppl.):1 - 60 (1999); Schena (ed.), Microarray Biochip: Tools and Technology, Eaton Publishing Company/BioTechniques Books Division (2000). Additionally, these nucleic acid microarrays include a substrate-bound plurality of nucleic acids in which the plurality of nucleic acids are disposed on a plurality of beads, rather than on a unitary planar substrate, as is described, *inter alia*, in Brenner *et al.*, *Proc. Natl. Acad. Sci. USA* 97(4):1665-1670 (2000). Examples of nucleic acid microarrays may be found in U.S. Patent Nos. 15 6,391,623, 6,383,754, 6,383,749, 6,380,377, 6,379,897, 6,376,191, 6,372,431, 6,351,712, 6,344,316, 6,316,193, 6,312,906, 6,309,828, 6,309,824, 6,306,643, 6,300,063, 6,287,850, 6,284,497, 6,284,465, 6,280,954, 6,262,216, 6,251,601, 6,245,518, 6,263,287, 6,251,601, 6,238,866, 6,228,575, 6,214,587, 6,203,989, 6,171,797, 6,103,474, 6,083,726, 6,054,274, 6,040,138, 6,083,726, 6,004,755, 6,001,309, 5,958,342, 5,952,180, 5,936,731, 5,843,655, 20 5,814,454, 5,837,196, 5,436,327, 5,412,087, and 5,405,783, the disclosures of which are incorporated herein by reference in their entireties.

In an alternative embodiment, a "microarray" may also refer to a "peptide microarray" or "protein microarray" having a substrate-bound collection or plurality of polypeptides, the binding to each of the plurality of bound polypeptides being separately detectable. Alternatively, the peptide microarray may have a plurality of binders, including but not limited to monoclonal antibodies, polyclonal antibodies, phage display binders, yeast 2 hybrid binders, and aptamers, which can specifically detect the binding of the polypeptides of this invention. The array may be based on autoantibody detection to the polypeptides of this invention, see Robinson *et al.*, *Nature Medicine* 8(3):295-301 25 30 (2002). Examples of peptide arrays may be found in WO 02/31463, WO 02/25288, WO 01/94946, WO 01/88162, WO 01/68671, WO 01/57259, WO 00/61806, WO 00/54046, WO 00/47774, WO 99/40434, WO 99/39210, and WO 97/42507 and U.S. Patent Nos.

6,268,210, 5,766,960, and 5,143,854, the disclosures of which are incorporated herein by reference in their entireties.

In addition, determination of the levels of the BSNA or BSP may be made in a multiplex manner using techniques described in WO 02/29109, WO 02/24959, WO

- 5 01/83502, WO01/73113, WO 01/59432, WO 01/57269, and WO 99/67641, the disclosures of which are incorporated herein by reference in their entireties.

The term “mutant”, “mutated”, or “mutation” when applied to nucleic acid sequences means that nucleotides in a nucleic acid sequence may be inserted, deleted or changed compared to a reference nucleic acid sequence. A single alteration may be made 10 at a locus (a point mutation) or multiple nucleotides may be inserted, deleted or changed at a single locus. In addition, one or more alterations may be made at any number of loci within a nucleic acid sequence. In a preferred embodiment of the present invention, the nucleic acid sequence is the wild type nucleic acid sequence encoding a BSP or is a BSNA. The nucleic acid sequence may be mutated by any method known in the art 15 including those mutagenesis techniques described *infra*.

The term “error-prone PCR” refers to a process for performing PCR under conditions where the copying fidelity of the DNA polymerase is low, such that a high rate of point mutations is obtained along the entire length of the PCR product. *See, e.g.,* Leung *et al.*, *Technique* 1: 11-15 (1989) and Caldwell *et al.*, *PCR Methods Applic.* 2: 28-33 20 (1992).

The term “oligonucleotide-directed mutagenesis” refers to a process which enables the generation of site-specific mutations in any cloned DNA segment of interest. *See, e.g.,* Reidhaar-Olson *et al.*, *Science* 241: 53-57 (1988).

The term “assembly PCR” refers to a process which involves the assembly of a 25 PCR product from a mixture of small DNA fragments. A large number of different PCR reactions occur in parallel in the same vial, with the products of one reaction priming the products of another reaction.

The term “sexual PCR mutagenesis” or “DNA shuffling” refers to a method of error-prone PCR coupled with forced homologous recombination between DNA 30 molecules of different but highly related DNA sequence *in vitro*, caused by random fragmentation of the DNA molecule based on sequence similarity, followed by fixation of the crossover by primer extension in an error-prone PCR reaction. *See, e.g.,* Stemmer,

Proc. Natl. Acad. Sci. U.S.A. 91: 10747-10751 (1994). DNA shuffling can be carried out between several related genes ("Family shuffling").

The term "in vivo mutagenesis" refers to a process of generating random mutations in any cloned DNA of interest which involves the propagation of the DNA in a strain of 5 bacteria such as *E. coli* that carries mutations in one or more of the DNA repair pathways. These "mutator" strains have a higher random mutation rate than that of a wild-type parent. Propagating the DNA in a mutator strain will eventually generate random mutations within the DNA.

The term "cassette mutagenesis" refers to any process for replacing a small region 10 of a double-stranded DNA molecule with a synthetic oligonucleotide "cassette" that differs from the native sequence. The oligonucleotide often contains completely and/or partially randomized native sequence.

The term "recursive ensemble mutagenesis" refers to an algorithm for protein engineering (protein mutagenesis) developed to produce diverse populations of 15 phenotypically related mutants whose members differ in amino acid sequence. This method uses a feedback mechanism to control successive rounds of combinatorial cassette mutagenesis. See, e.g., Arkin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89: 7811-7815 (1992).

The term "exponential ensemble mutagenesis" refers to a process for generating 20 combinatorial libraries with a high percentage of unique and functional mutants, wherein small groups of residues are randomized in parallel to identify, at each altered position, amino acids which lead to functional proteins. See, e.g., Delegrave *et al.*, *Biotechnology Research* 11: 1548-1552 (1993); Arnold, *Current Opinion in Biotechnology* 4: 450-455 (1993).

"Operatively linked" expression control sequences refers to a linkage in which the 25 expression control sequence is either contiguous with the gene of interest to control the gene of interest, or acts in *trans* or at a distance to control the gene of interest.

The term "expression control sequence" as used herein refers to polynucleotide sequences which are necessary to affect the expression of coding sequences to which they are operatively linked. Expression control sequences are sequences which control the 30 transcription, post-transcriptional events and translation of nucleic acid sequences. Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation signals; sequences that stabilize cytoplasmic mRNA; sequences that

enhance translation efficiency (*e.g.*, ribosome binding sites); sequences that enhance protein stability; and when desired, sequences that enhance protein secretion. The nature of such control sequences differs depending upon the host organism; in prokaryotes, such control sequences generally include promoter, ribosomal binding site, and transcription 5 termination sequence. The term “control sequences” is intended to include, at a minimum, all components whose presence is essential for expression, and can also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences.

The term “vector,” as used herein, is intended to refer to a nucleic acid molecule 10 capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a circular double-stranded DNA loop into which additional DNA segments may be ligated. Other vectors include cosmids, bacterial artificial chromosomes (BAC) and yeast artificial chromosomes (YAC). Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral 15 genome. Viral vectors that infect bacterial cells are referred to as bacteriophages. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of 20 directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “recombinant expression vectors” (or simply, “expression vectors”). In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, “plasmid” and “vector” may be used interchangeably as the plasmid is the most commonly used form of vector. However, the 25 invention is intended to include other forms of expression vectors that serve equivalent functions.

The term “recombinant host cell” (or simply “host cell”), as used herein, is intended to refer to a cell into which a recombinant expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular 30 subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term “host cell” as used herein.

As used herein, the phrase "open reading frame" and the equivalent acronym "ORF" refers to that portion of a transcript-derived nucleic acid that can be translated in its entirety into a sequence of contiguous amino acids. As so defined, an ORF has length, measured in nucleotides, exactly divisible by 3. As so defined, an ORF need not encode 5 the entirety of a natural protein.

As used herein, the phrase "ORF-encoded peptide" refers to the predicted or actual translation of an ORF.

As used herein, the phrase "degenerate variant" of a reference nucleic acid sequence is meant to be inclusive of all nucleic acid sequences that can be directly 10 translated, using the standard genetic code, to provide an amino acid sequence identical to that translated from the reference nucleic acid sequence.

The term "polypeptide" encompasses both naturally occurring and non-naturally occurring proteins and polypeptides, as well as polypeptide fragments and polypeptide mutants, derivatives and analogs thereof. A polypeptide may be monomeric or polymeric. 15 Further, a polypeptide may comprise a number of different modules within a single polypeptide each of which has one or more distinct activities. A preferred polypeptide in accordance with the invention comprises a BSP encoded by a nucleic acid molecule of the instant invention, or a fragment, mutant, analog or derivative thereof.

The term "isolated protein" or "isolated polypeptide" is a protein or polypeptide 20 that by virtue of its origin or source of derivation (1) is not associated with naturally associated components that accompany it in its native state, (2) is free of other proteins from the same species (3) is expressed by a cell from a different species, or (4) does not occur in nature. Thus, a polypeptide that is chemically synthesized or synthesized in a cellular system different from the cell from which it naturally originates will be "isolated" 25 from its naturally associated components. A polypeptide or protein may also be rendered substantially free of naturally associated components by isolation, using protein purification techniques well known in the art.

A protein or polypeptide is "substantially pure," "substantially homogeneous" or "substantially purified" when at least about 60% to 75% of a sample exhibits a single 30 species of polypeptide. The polypeptide or protein may be monomeric or multimeric. A substantially pure polypeptide or protein will typically comprise about 50%, 60%, 70%, 80% or 90% W/W of a protein sample, more usually about 95%, and preferably will be over 99% pure. Protein purity or homogeneity may be determined by a number of means

well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band upon staining the gel with a stain well known in the art. For certain purposes, higher resolution may be provided by using HPLC or other means well known in the art for purification.

5 The term "fragment" when used herein with respect to polypeptides of the present invention refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion compared to a full-length BSP. In a preferred embodiment, the fragment is a contiguous sequence in which the amino acid sequence of the fragment is identical to the corresponding positions in the naturally occurring polypeptide. Fragments typically are at
10 least 5, 6, 7, 8, 9 or 10 amino acids long, preferably at least 12, 14, 16 or 18 amino acids long, more preferably at least 20 amino acids long, more preferably at least 25, 30, 35, 40 or 45, amino acids, even more preferably at least 50 or 60 amino acids long, and even more preferably at least 70 amino acids long.

A "derivative" when used herein with respect to polypeptides of the present
15 invention refers to a polypeptide which is substantially similar in primary structural sequence to a BSP but which includes, *e.g.*, *in vivo* or *in vitro* chemical and biochemical modifications that are not found in the BSP. Such modifications include, for example, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide
20 derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristylation, oxidation, proteolytic processing,
25 phosphorylation, prenylation, racemization, selenylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. Other modifications include, *e.g.*, labeling with radionuclides, and various enzymatic modifications, as will be readily appreciated by those skilled in the art. A variety of methods for labeling polypeptides and of substituents or labels useful for such purposes
30 are well known in the art, and include radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , ^{14}C and ^3H , ligands which bind to labeled antiligands (*e.g.*, antibodies), fluorophores, chemiluminescent agents, enzymes, and antiligands which can serve as specific binding pair members for a labeled ligand. The choice of label depends on the sensitivity required,

ease of conjugation with the primer, stability requirements, and available instrumentation. Methods for labeling polypeptides are well known in the art. See Ausubel (1992), *supra*; Ausubel (1999), *supra*.

The term "fusion protein" refers to polypeptides of the present invention coupled 5 to a heterologous amino acid sequence. Fusion proteins are useful because they can be constructed to contain two or more desired functional elements from two or more different proteins. A fusion protein comprises at least 10 contiguous amino acids from a polypeptide of interest, more preferably at least 20 or 30 amino acids, even more preferably at least 40, 50 or 60 amino acids, yet more preferably at least 75, 100 or 125 10 amino acids. Fusion proteins can be produced recombinantly by constructing a nucleic acid sequence that encodes the polypeptide or a fragment thereof in frame with a nucleic acid sequence encoding a different protein or peptide and then expressing the fusion protein. Alternatively, a fusion protein can be produced chemically by crosslinking the polypeptide or a fragment thereof to another protein.

15 The term "analog" refers to both polypeptide analogs and non-peptide analogs. The term "polypeptide analog" as used herein refers to a polypeptide that is comprised of a segment of at least 25 amino acids that has substantial identity to a portion of an amino acid sequence but which contains non-natural amino acids or non-natural inter-residue bonds. In a preferred embodiment, the analog has the same or similar biological activity 20 as the native polypeptide. Typically, polypeptide analogs comprise a conservative amino acid substitution (or insertion or deletion) with respect to the naturally occurring sequence. Analogs typically are at least 20 amino acids long, preferably at least 50 amino acids long or longer, and can often be as long as a full-length naturally occurring polypeptide.

The term "non-peptide analog" refers to a compound with properties that are 25 analogous to those of a reference polypeptide. A non-peptide compound may also be termed a "peptide mimetic" or a "peptidomimetic." Such compounds are often developed with the aid of computerized molecular modeling. Peptide mimetics that are structurally similar to useful peptides may be used to produce an equivalent effect. Generally, peptidomimetics are structurally similar to a paradigm polypeptide (*i.e.*, a polypeptide that 30 has a desired biochemical property or pharmacological activity), but have one or more peptide linkages optionally replaced by a linkage selected from the group consisting of: --CH₂NH--, --CH₂S--, --CH₂-CH₂--, --CH=CH--(cis and trans), --COCH₂--, --CH(OH)CH₂--, and --CH₂SO--, by methods well known in the art. Systematic

substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (*e.g.*, D-lysine in place of L-lysine) may also be used to generate more stable peptides. In addition, constrained peptides comprising a consensus sequence or a substantially identical consensus sequence variation may be generated by methods known 5 in the art (Rizo *et al.*, *Ann. Rev. Biochem.* 61:387-418 (1992)). For example, one may add internal cysteine residues capable of forming intramolecular disulfide bridges which cyclize the peptide.

The term "mutant" or "mutein" when referring to a polypeptide of the present invention relates to an amino acid sequence containing substitutions, insertions or 10 deletions of one or more amino acids compared to the amino acid sequence of a BSP. A mutein may have one or more amino acid point substitutions, in which a single amino acid at a position has been changed to another amino acid, one or more insertions and/or deletions, in which one or more amino acids are inserted or deleted, respectively, in the sequence of the naturally occurring protein, and/or truncations of the amino acid sequence 15 at either or both the amino or carboxy termini. Further, a mutein may have the same or different biological activity as the naturally occurring protein. For instance, a mutein may have an increased or decreased biological activity. A mutein has at least 50% sequence similarity to the wild type protein, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are muteins having 80%, 85% or 90% 20 sequence similarity to a BSP. In an even more preferred embodiment, a mutein exhibits 95% sequence identity, even more preferably 97%, even more preferably 98% and even more preferably 99%. Sequence similarity may be measured by any common sequence analysis algorithm, such as GAP or BESTFIT or other variation Smith-Waterman alignment. *See*, T. F. Smith and M. S. Waterman, *J. Mol. Biol.* 147:195-197 (1981) and 25 W.R. Pearson, *Genomics* 11:635-650 (1991).

Preferred amino acid substitutions are those which: (1) reduce susceptibility to proteolysis, (2) reduce susceptibility to oxidation, (3) alter binding affinity for forming protein complexes, (4) alter binding affinity or enzymatic activity, and (5) confer or modify other physicochemical or functional properties of such analogs. For example, 30 single or multiple amino acid substitutions (preferably conservative amino acid substitutions) may be made in the naturally occurring sequence (preferably in the portion of the polypeptide outside the domain(s) forming intermolecular contacts. In a preferred embodiment, the amino acid substitutions are moderately conservative substitutions or

conservative substitutions. In a more preferred embodiment, the amino acid substitutions are conservative substitutions. A conservative amino acid substitution should not substantially change the structural characteristics of the parent sequence (e.g., a replacement amino acid should not tend to disrupt a helix that occurs in the parent sequence, or disrupt other types of secondary structure that characterize the parent sequence). Examples of art-recognized polypeptide secondary and tertiary structures are described in Creighton (ed.), Proteins, Structures and Molecular Principles, W. H. Freeman and Company (1984); Branden *et al.* (ed.), Introduction to Protein Structure, Garland Publishing (1991); Thornton *et al.*, *Nature* 354:105-106 (1991).

As used herein, the twenty conventional amino acids and their abbreviations follow conventional usage. See Golub *et al.* (eds.), Immunology - A Synthesis 2nd Ed., Sinauer Associates (1991). Stereoisomers (e.g., D-amino acids) of the twenty conventional amino acids, unnatural amino acids such as α -, α -disubstituted amino acids, N-alkyl amino acids, and other unconventional amino acids may also be suitable components for polypeptides of the present invention. Examples of unconventional amino acids include: 4-hydroxyproline, γ -carboxyglutamate, ϵ -N,N,N-trimethyllysine, ϵ -N-acetyllysine, O-phosphoserine, N-acetylserine, N-formylmethionine, 3-methylhistidine, 5-hydroxylysine, s-N-methylarginine, and other similar amino acids and imino acids (e.g., 4-hydroxyproline). In the polypeptide notation used herein, the lefthand direction is the amino terminal direction and the right hand direction is the carboxy-terminal direction, in accordance with standard usage and convention.

By "homology" or "homologous" when referring to a polypeptide of the present invention it is meant polypeptides from different organisms with a similar sequence to the encoded amino acid sequence of a BSP and a similar biological activity or function. Although two polypeptides are said to be "homologous," this does not imply that there is necessarily an evolutionary relationship between the polypeptides. Instead, the term "homologous" is defined to mean that the two polypeptides have similar amino acid sequences and similar biological activities or functions. In a preferred embodiment, a homologous polypeptide is one that exhibits 50% sequence similarity to BSP, preferred is 60% sequence similarity, more preferred is 70% sequence similarity. Even more preferred are homologous polypeptides that exhibit 80%, 85% or 90% sequence similarity to a BSP. In yet a more preferred embodiment, a homologous polypeptide exhibits 95%, 97%, 98% or 99% sequence similarity.

When "sequence similarity" is used in reference to polypeptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions. In a preferred embodiment, a polypeptide that has "sequence similarity" comprises conservative or moderately conservative amino acid substitutions. A 5 "conservative amino acid substitution" is one in which an amino acid residue is substituted by another amino acid residue having a side chain (R group) with similar chemical properties (e.g., charge or hydrophobicity). In general, a conservative amino acid substitution will not substantially change the functional properties of a protein. In cases where two or more amino acid sequences differ from each other by conservative 10 substitutions, the percent sequence identity or degree of similarity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. See, e.g., Pearson, *Methods Mol. Biol.* 24: 307-31 (1994).

For instance, the following six groups each contain amino acids that are 15 conservative substitutions for one another:

- 1) Serine (S), Threonine (T);
- 2) Aspartic Acid (D), Glutamic Acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 20 5) Isoleucine (I), Leucine (L), Methionine (M), Alanine (A), Valine (V), and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Alternatively, a conservative replacement is any change having a positive value in the PAM250 log-likelihood matrix disclosed in Gonnet *et al.*, *Science* 256: 1443-45 (1992). A "moderately conservative" replacement is any change having a nonnegative 25 value in the PAM250 log-likelihood matrix.

Sequence similarity for polypeptides, which is also referred to as sequence identity, is typically measured using sequence analysis software. Protein analysis software matches similar sequences using measures of similarity assigned to various substitutions, deletions and other modifications, including conservative amino acid substitutions. For 30 instance, GCG contains programs such as "Gap" and "Bestfit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of

organisms or between a wild type protein and a mutein thereof. *See, e.g.*, GCG Version 6.1. Other programs include FASTA, discussed *supra*.

A preferred algorithm when comparing a sequence of the invention to a database containing a large number of sequences from different organisms is the computer program

- 5 BLAST, especially blastp or tblastn. *See, e.g.*, Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990); Altschul *et al.*, *Nucleic Acids Res.* 25:3389-402 (1997). Preferred parameters for blastp are:

Expectation value: 10 (default)
Filter: seg (default)
10 Cost to open a gap: 11 (default)
Cost to extend a gap: 1 (default)
Max. alignments: 100 (default)
Word size: 11 (default)
No. of descriptions: 100 (default)
15 Penalty Matrix: BLOSUM62

The length of polypeptide sequences compared for homology will generally be at least about 16 amino acid residues, usually at least about 20 residues, more usually at least about 24 residues, typically at least about 28 residues, and preferably more than about 35 residues. When searching a database containing sequences from a large number of
20 different organisms, it is preferable to compare amino acid sequences.

Algorithms other than blastp for database searching using amino acid sequences are known in the art. For instance, polypeptide sequences can be compared using FASTA, a program in GCG Version 6.1. FASTA (*e.g.*, FASTA2 and FASTA3) provides alignments and percent sequence identity of the regions of the best overlap between the
25 query and search sequences (Pearson (1990), *supra*; Pearson (2000), *supra*). For example, percent sequence identity between amino acid sequences can be determined using FASTA with its default or recommended parameters (a word size of 2 and the PAM250 scoring matrix), as provided in GCG Version 6.1.

An “antibody” refers to an intact immunoglobulin, or to an antigen-binding portion
30 thereof that competes with the intact antibody for specific binding to a molecular species, *e.g.*, a polypeptide of the instant invention. Antigen-binding portions may be produced by recombinant DNA techniques or by enzymatic or chemical cleavage of intact antibodies.

Antigen-binding portions include, *inter alia*, Fab, Fab', F(ab')₂, Fv, dAb, and

complementarity determining region (CDR) fragments, single-chain antibodies (scFv), chimeric antibodies, diabodies and polypeptides that contain at least a portion of an immunoglobulin that is sufficient to confer specific antigen binding to the polypeptide. A Fab fragment is a monovalent fragment consisting of the VL, VH, CL and CH1 domains; a F(ab')₂ fragment is a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; a Fd fragment consists of the VH and CH1 domains; a Fv fragment consists of the VL and VH domains of a single arm of an antibody; and a dAb fragment consists of a VH domain. *See, e.g., Ward et al., Nature 341: 544-546 (1989).*

By "bind specifically" and "specific binding" as used herein it is meant the ability of the antibody to bind to a first molecular species in preference to binding to other molecular species with which the antibody and first molecular species are admixed. An antibody is said to "recognize" a first molecular species when it can bind specifically to that first molecular species.

A single-chain antibody (scFv) is an antibody in which VL and VH regions are paired to form a monovalent molecule via a synthetic linker that enables them to be made as a single protein chain. *See, e.g., Bird et al., Science 242: 423-426 (1988); Huston et al., Proc. Natl. Acad. Sci. USA 85: 5879-5883 (1988).* Diabodies are bivalent, bispecific antibodies in which VH and VL domains are expressed on a single polypeptide chain, but using a linker that is too short to allow for pairing between the two domains on the same chain, thereby forcing the domains to pair with complementary domains of another chain and creating two antigen binding sites. *See e.g., Holliger et al., Proc. Natl. Acad. Sci. USA 90: 6444-6448 (1993); Poljak et al., Structure 2: 1121-1123 (1994).* One or more CDRs may be incorporated into a molecule either covalently or noncovalently to make it an immunoadhesin. An immunoadhesin may incorporate the CDR(s) as part of a larger polypeptide chain, may covalently link the CDR(s) to another polypeptide chain, or may incorporate the CDR(s) noncovalently. The CDRs permit the immunoadhesin to specifically bind to a particular antigen of interest. A chimeric antibody is an antibody that contains one or more regions from one antibody and one or more regions from one or more other antibodies.

An antibody may have one or more binding sites. If there is more than one binding site, the binding sites may be identical to one another or may be different. For instance, a naturally occurring immunoglobulin has two identical binding sites, a single-chain

antibody or Fab fragment has one binding site, while a “bispecific” or “bifunctional” antibody has two different binding sites.

An “isolated antibody” is an antibody that (1) is not associated with naturally-associated components, including other naturally-associated antibodies, that accompany it 5 in its native state, (2) is free of other proteins from the same species, (3) is expressed by a cell from a different species, or (4) does not occur in nature. It is known that purified proteins, including purified antibodies, may be stabilized with non-naturally-associated components. The non-naturally-associated component may be a protein, such as albumin (e.g., BSA) or a chemical such as polyethylene glycol (PEG).

10 A “neutralizing antibody” or “an inhibitory antibody” is an antibody that inhibits the activity of a polypeptide or blocks the binding of a polypeptide to a ligand that normally binds to it. An “activating antibody” is an antibody that increases the activity of a polypeptide.

15 The term “epitope” includes any protein determinant capable of specific binding to an immunoglobulin or T-cell receptor. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three-dimensional structural characteristics, as well as specific charge characteristics. An antibody is said to specifically bind an antigen when the dissociation constant is less than 1 μ M, preferably less than 100 nM and most preferably 20 less than 10 nM.

The term “patient” includes human and veterinary subjects.

Throughout this specification and claims, the word “comprise,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

25 The term “breast specific” refers to a nucleic acid molecule or polypeptide that is expressed predominantly in the breast as compared to other tissues in the body. In a preferred embodiment, a “breast specific” nucleic acid molecule or polypeptide is detected at a level that is 1.5-fold higher than any other tissue in the body. In a more preferred embodiment, the “breast specific” nucleic acid molecule or polypeptide is detected at a 30 level that is 2-fold higher than any other tissue in the body, more preferably 5-fold higher, still more preferably at least 10-fold, 15-fold, 20-fold, 25-fold, 50-fold or 100-fold higher than any other tissue in the body. Nucleic acid molecule levels may be measured by nucleic acid hybridization, such as Northern blot hybridization, or quantitative PCR.

Polypeptide levels may be measured by any method known to accurately quantitate protein levels, such as Western blot analysis.

Nucleic Acid Molecules, Regulatory Sequences, Vectors, Host Cells and Recombinant Methods of Making Polypeptides

5 *Nucleic Acid Molecules*

One aspect of the invention provides isolated nucleic acid molecules that are specific to the breast or to breast cells or tissue or that are derived from such nucleic acid molecules. These isolated breast specific nucleic acids (BSNAs) may comprise cDNA genomic DNA, RNA, or a combination thereof, a fragment of one of these nucleic acids, 10 or may be a non-naturally occurring nucleic acid molecule. A BSNA may be derived from an animal. In a preferred embodiment, the BSNA is derived from a human or other mammal. In a more preferred embodiment, the BSNA is derived from a human or other primate. In an even more preferred embodiment, the BSNA is derived from a human.

In a preferred embodiment, the nucleic acid molecule encodes a polypeptide that 15 is specific to breast, a breast-specific polypeptide (BSP). In a more preferred embodiment, the nucleic acid molecule encodes a polypeptide that comprises an amino acid sequence of SEQ ID NO: 96-232. In another highly preferred embodiment, the nucleic acid molecule comprises a nucleic acid sequence of SEQ ID NO: 1-95. Nucleotide sequences of the instantly-described nucleic acid molecules were determined by assembling several DNA 20 molecules from either public or proprietary databases. Some of the underlying DNA sequences are the result, directly or indirectly, of at least one enzymatic polymerization reaction (*e.g.*, reverse transcription and/or polymerase chain reaction) using an automated sequencer (such as the MegaBACE™ 1000, Amersham Biosciences, Sunnyvale, CA, USA).

25 Nucleic acid molecules of the present invention may also comprise sequences that selectively hybridize to a nucleic acid molecule encoding a BSNA or a complement or antisense thereof. The hybridizing nucleic acid molecule may or may not encode a polypeptide or may or may not encode a BSP. However, in a preferred embodiment, the hybridizing nucleic acid molecule encodes a BSP. In a more preferred embodiment, the 30 invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence of SEQ ID NO: 96-232. In an even more preferred

embodiment, the invention provides a nucleic acid molecule that selectively hybridizes to a nucleic acid molecule comprising the nucleic acid sequence of SEQ ID NO: 1-95 or the antisense sequence thereof. Preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a

5 BSP under low stringency conditions. More preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a BSP under moderate stringency conditions. Most preferably, the nucleic acid molecule selectively hybridizes to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a BSP under high stringency conditions. In

10 a preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense sequence of a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence of SEQ ID NO: 96-232. In a more preferred embodiment, the nucleic acid molecule hybridizes under low, moderate or high stringency conditions to a nucleic acid molecule or the antisense

15 sequence of a nucleic acid molecule comprising a nucleic acid sequence selected from SEQ ID NO: 1-95.

Nucleic acid molecules of the present invention may also comprise nucleic acid sequences that exhibit substantial sequence similarity to a nucleic acid encoding a BSP or a complement of the encoding nucleic acid molecule. In this embodiment, it is preferred

20 that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule encoding human BSP. More preferred is a nucleic acid molecule exhibiting substantial sequence similarity to a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 96-232. By substantial sequence similarity it is meant a nucleic acid molecule having at least 60%, more preferably at least 70%, even

25 more preferably at least 80% and even more preferably at least 85% sequence identity with a nucleic acid molecule encoding a BSP, such as a polypeptide having an amino acid sequence of SEQ ID NO: 96-232. In a more preferred embodiment, the similar nucleic acid molecule is one that has at least 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at least 98%, and still more preferably at least 99%

30 sequence identity with a nucleic acid molecule encoding a BSP. Most preferred in this embodiment is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a nucleic acid molecule encoding a BSP.

The nucleic acid molecules of the present invention are also inclusive of those exhibiting substantial sequence similarity to a BSNA or its complement. In this embodiment, it is preferred that the nucleic acid molecule exhibit substantial sequence similarity to a nucleic acid molecule having a nucleic acid sequence of SEQ ID NO: 1-95.

- 5 By substantial sequence similarity it is meant a nucleic acid molecule that has at least 60%, more preferably at least 70%, even more preferably at least 80% and even more preferably at least 85% sequence identity with a BSNA, such as one having a nucleic acid sequence of SEQ ID NO: 1-95. More preferred is a nucleic acid molecule that has at least 90%, more preferably at least 95%, more preferably at least 97%, even more preferably at 10 least 98%, and still more preferably at least 99% sequence identity with a BSNA. Most preferred is a nucleic acid molecule that has at least 99.5%, 99.6%, 99.7%, 99.8% or 99.9% sequence identity with a BSNA.

Nucleic acid molecules that exhibit substantial sequence similarity are inclusive of sequences that exhibit sequence identity over their entire length to a BSNA or to a nucleic acid molecule encoding a BSP, as well as sequences that are similar over only a part of its length. In this case, the part is at least 50 nucleotides of the BSNA or the nucleic acid molecule encoding a BSP, preferably at least 100 nucleotides, more preferably at least 150 or 200 nucleotides, even more preferably at least 250 or 300 nucleotides, still more preferably at least 400 or 500 nucleotides.

- 15 20 The substantially similar nucleic acid molecule may be a naturally occurring one that is derived from another species, especially one derived from another primate, wherein the similar nucleic acid molecule encodes an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 96-232 or demonstrates significant sequence identity to the nucleotide sequence of SEQ ID NO: 1-95. The similar nucleic acid 25 molecule may also be a naturally occurring nucleic acid molecule from a human, when the BSNA is a member of a gene family. The similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-primate, mammalian species, including without limitation, domesticated species, *e.g.*, dog, cat, mouse, rat, rabbit, hamster, cow, horse and pig; and wild animals, *e.g.*, monkey, fox, lions, tigers, 30 bears, giraffes, zebras, etc. The substantially similar nucleic acid molecule may also be a naturally occurring nucleic acid molecule derived from a non-mammalian species, such as birds or reptiles. The naturally occurring substantially similar nucleic acid molecule may be isolated directly from humans or other species. In another embodiment, the

substantially similar nucleic acid molecule may be one that is experimentally produced by random mutation of a nucleic acid molecule. In another embodiment, the substantially similar nucleic acid molecule may be one that is experimentally produced by directed mutation of a BSNA. In a preferred embodiment, the substantially similar nucleic acid molecule is a BSNA.

The nucleic acid molecules of the present invention are also inclusive of allelic variants of a BSNA or a nucleic acid encoding a BSP. For example, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes and the sequence determined from one individual of a species may differ from other allelic forms present within the population. More than 1.4 million SNPs have already been identified in the human genome, International Human Genome Sequencing Consortium, *Nature* 409: 860-921 (2001) – Variants with small deletions and insertions of more than a single nucleotide are also found in the general population, and often do not alter the function of the protein. In addition, amino acid substitutions occur frequently among natural allelic variants, and often do not substantially change protein function.

In a preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that encodes a BSP. In a more preferred embodiment, the gene is transcribed into a mRNA that encodes a BSP comprising an amino acid sequence of SEQ ID NO: 96-232. In another preferred embodiment, the allelic variant is a variant of a gene, wherein the gene is transcribed into a mRNA that is a BSNA. In a more preferred embodiment, the gene is transcribed into a mRNA that comprises the nucleic acid sequence of SEQ ID NO: 1-95. Also preferred is that the allelic variant be a naturally occurring allelic variant in the species of interest, particularly human.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences comprising a part of a nucleic acid sequence of the instant invention. The part may or may not encode a polypeptide, and may or may not encode a polypeptide that is a BSP. In a preferred embodiment, the part encodes a BSP. In one embodiment, the nucleic acid molecule comprises a part of a BSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that hybridizes or exhibits substantial sequence similarity to a BSNA. In another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that is an allelic variant of a BSNA. In yet another embodiment, the nucleic acid molecule comprises a part of a nucleic acid molecule that encodes a BSP. A part comprises at least 10 nucleotides, more preferably at

least 15, 17, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400 or 500 nucleotides. The maximum size of a nucleic acid part is one nucleotide shorter than the sequence of the nucleic acid molecule encoding the full-length protein.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences that encode fusion proteins, homologous proteins, polypeptide fragments, muteins and polypeptide analogs, as described *infra*.

Nucleic acid molecules of the present invention are also inclusive of nucleic acid sequences containing modifications of the native nucleic acid molecule. Examples of such modifications include, but are not limited to, nonnative internucleoside bonds, post-synthetic modifications or altered nucleotide analogues. One having ordinary skill in the art would recognize that the type of modification that may be made will depend upon the intended use of the nucleic acid molecule. For instance, when the nucleic acid molecule is used as a hybridization probe, the range of such modifications will be limited to those that permit sequence-discriminating base pairing of the resulting nucleic acid. When used to direct expression of RNA or protein *in vitro* or *in vivo*, the range of such modifications will be limited to those that permit the nucleic acid to function properly as a polymerization substrate. When the isolated nucleic acid is used as a therapeutic agent, the modifications will be limited to those that do not confer toxicity upon the isolated nucleic acid.

Accordingly, in one embodiment, a nucleic acid molecule may include nucleotide analogues that incorporate labels that are directly detectable, such as radiolabels or fluorophores, or nucleotide analogues that incorporate labels that can be visualized in a subsequent reaction, such as biotin or various haptens. The labeled nucleic acid molecules are particularly useful as hybridization probes.

Common radiolabeled analogues include those labeled with ^{33}P , ^{32}P , and ^{35}S , such as α - ^{32}P -dATP, α - ^{32}P -dCTP, α - ^{32}P -dGTP, α - ^{32}P -dTTP, α - ^{32}P -3'dATP, α - ^{32}P -ATP, α - ^{32}P -CTP, α - ^{32}P -GTP, α - ^{32}P -UTP, α - ^{35}S -dATP, γ - ^{35}S -GTP, γ - ^{33}P -dATP, and the like.

Commercially available fluorescent nucleotide analogues readily incorporated into the nucleic acids of the present invention include Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy3-dUTP (Amersham Biosciences, Piscataway, New Jersey, USA), fluorescein-12-dUTP, tetramethylrhodamine-6-dUTP, Texas Red®-5-dUTP, Cascade Blue®-7-dUTP, BODIPY® FL-14-dUTP, BODIPY® TMR-14-dUTP, BODIPY® TR-14-dUTP, Rhodamine Green™-5-dUTP, Oregon Green® 488-5-dUTP, Texas Red®-12-dUTP,

BODIPY® 630/650-14-dUTP, BODIPY® 650/665-14-dUTP, Alexa Fluor® 488-5-dUTP, Alexa Fluor® 532-5-dUTP, Alexa Fluor® 568-5-dUTP, Alexa Fluor® 594-5-dUTP, Alexa Fluor® 546-14-dUTP, fluorescein-12-UTP, tetramethylrhodamine-6-UTP, Texas

Red®-5-UTP, Cascade Blue®-7-UTP, BODIPY® FL-14-UTP, BODIPY® TMR-14-UTP,

- 5 BODIPY® TR-14-UTP, Rhodamine Green™-5-UTP, Alexa Fluor® 488-5-UTP, Alexa Fluor® 546-14-UTP (Molecular Probes, Inc. Eugene, OR, USA). One may also custom synthesize nucleotides having other fluorophores. *See Henegariu et al., Nature Biotechnol.* 18: 345-348 (2000).

Haptens that are commonly conjugated to nucleotides for subsequent labeling

- 10 include biotin (biotin-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA; biotin-21-UTP, biotin-21-dUTP, Clontech Laboratories, Inc., Palo Alto, CA, USA), digoxigenin (DIG-11-dUTP, alkali labile, DIG-11-UTP, Roche Diagnostics Corp., Indianapolis, IN, USA), and dinitrophenyl (dinitrophenyl-11-dUTP, Molecular Probes, Inc., Eugene, OR, USA).

- 15 Nucleic acid molecules of the present invention can be labeled by incorporation of labeled nucleotide analogues into the nucleic acid. Such analogues can be incorporated by enzymatic polymerization, such as by nick translation, random priming, polymerase chain reaction (PCR), terminal transferase tailing, and end-filling of overhangs, for DNA molecules, and *in vitro* transcription driven, e.g., from phage promoters, such as T7, T3, 20 and SP6, for RNA molecules. Commercial kits are readily available for each such labeling approach. Analogues can also be incorporated during automated solid phase chemical synthesis. Labels can also be incorporated after nucleic acid synthesis, with the 5' phosphate and 3' hydroxyl providing convenient sites for post-synthetic covalent attachment of detectable labels.

- 25 Other post-synthetic approaches also permit internal labeling of nucleic acids. For example, fluorophores can be attached using a cisplatin reagent that reacts with the N7 of guanine residues (and, to a lesser extent, adenine bases) in DNA, RNA, and Peptide Nucleic Acids (PNA) to provide a stable coordination complex between the nucleic acid and fluorophore label (Universal Linkage System) (available from Molecular Probes, Inc., 30 Eugene, OR, USA and Amersham Pharmacia Biotech, Piscataway, NJ, USA); *see Alers et al., Genes, Chromosomes & Cancer* 25: 301- 305 (1999); Jelsma et al., *J. NIH Res.* 5: 82 (1994); Van Belkum et al., *BioTechniques* 16: 148-153 (1994). Alternatively, nucleic acids can be labeled using a disulfide-containing linker (FastTag™ Reagent, Vector

Laboratories, Inc., Burlingame, CA, USA) that is photo- or thermally coupled to the target nucleic acid using aryl azide chemistry; after reduction, a free thiol is available for coupling to a hapten, fluorophore, sugar, affinity ligand, or other marker.

- One or more independent or interacting labels can be incorporated into the nucleic acid molecules of the present invention. For example, both a fluorophore and a moiety that in proximity thereto acts to quench fluorescence can be included to report specific hybridization through release of fluorescence quenching or to report exonucleotidic excision. *See, e.g.*, Tyagi *et al.*, *Nature Biotechnol.* 14: 303-308 (1996); Tyagi *et al.*, *Nature Biotechnol.* 16: 49-53 (1998); Sokol *et al.*, *Proc. Natl. Acad. Sci. USA* 95: 11538-11543 (1998); Kostrikis *et al.*, *Science* 279: 1228-1229 (1998); Marras *et al.*, *Genet. Anal.* 14: 151-156 (1999); Holland *et al.*, *Proc. Natl. Acad. Sci. USA* 88: 7276-7280 (1991); Heid *et al.*, *Genome Res.* 6(10): 986-94 (1996); Kuimelis *et al.*, *Nucleic Acids Symp. Ser.* (37): 255-6 (1997); and U.S. Patent Nos. 5,846,726, 5,925,517, 5,925,517, 5,723,591 and 5,538,848, the disclosures of which are incorporated herein by reference in their entireties.

- Nucleic acid molecules of the present invention may also be modified by altering one or more native phosphodiester internucleoside bonds to more nuclease-resistant, internucleoside bonds. *See* Hartmann *et al.* (eds.), Manual of Antisense Methodology: Perspectives in Antisense Science, Kluwer Law International (1999); Stein *et al.* (eds.), Applied Antisense Oligonucleotide Technology, Wiley-Liss (1998); Chadwick *et al.* (eds.), Oligonucleotides as Therapeutic Agents – Symposium No. 209, John Wiley & Son Ltd (1997). Such altered internucleoside bonds are often desired for techniques or for targeted gene correction, Gamper *et al.*, *Nucl. Acids Res.* 28(21): 4332-4339 (2000). For double-stranded RNA inhibition which may utilize either natural ds RNA or ds RNA modified in its, sugar, phosphate or base, *see* Hannon, *Nature* 418(11): 244-251 (2002); Fire *et al.* in WO 99/32619; Tuschl *et al.* in US2002/0086356; Krueutzer *et al.* in WO 00/44895, the disclosures of which are incorporated herein by reference in their entirety. For circular antisense, *see* Kool in U.S. Patent No. 5,426,180, the disclosure of which is incorporated herein by reference in its entirety.

- Modified oligonucleotide backbones include, without limitation, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including

- 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'.
- 5 Representative U.S. Patents that teach the preparation of the above phosphorus-containing linkages include, but are not limited to, U.S. Patent Nos. 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050, the disclosures of which are incorporated herein by reference in their entireties. In a preferred embodiment, the modified internucleoside linkages may be used for antisense techniques.

Other modified oligonucleotide backbones do not include a phosphorus atom, but have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or 15 more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino 20 backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂ component parts. Representative U.S. patents that teach the preparation of the above backbones include, but are not limited to, U.S. Patent Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 25 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437 and 5,677,439; the disclosures of which are incorporated herein by reference in their entireties.

In other preferred nucleic acid molecules, both the sugar and the internucleoside linkage are replaced with novel groups, such as peptide nucleic acids (PNA). In PNA 30 compounds, the phosphodiester backbone of the nucleic acid is replaced with an amide-containing backbone, in particular by repeating N-(2-aminoethyl) glycine units linked by amide bonds. Nucleobases are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone, typically by methylene carbonyl linkages. PNA can be

synthesized using a modified peptide synthesis protocol. PNA oligomers can be synthesized by both Fmoc and tBoc methods. Representative U.S. patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Patent Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference in its entirety. Automated PNA synthesis is readily achievable on commercial synthesizers (see, e.g., "PNA User's Guide," Rev. 2, February 1998, Perseptive Biosystems Part No. 60138, Applied Biosystems, Inc., Foster City, CA). PNA molecules are advantageous for a number of reasons. First, because the PNA backbone is uncharged, PNA/DNA and PNA/RNA duplexes have a higher thermal stability than is found in DNA/DNA and DNA/RNA duplexes. The T_m of a PNA/DNA or PNA/RNA duplex is generally 1°C higher per base pair than the T_m of the corresponding DNA/DNA or DNA/RNA duplex (in 100 mM NaCl). Second, PNA molecules can also form stable PNA/DNA complexes at low ionic strength, under conditions in which DNA/DNA duplex formation does not occur. Third, PNA also demonstrates greater specificity in binding to complementary DNA because a PNA/DNA mismatch is more destabilizing than DNA/DNA mismatch. A single mismatch in mixed a PNA/DNA 15-mer lowers the T_m by 8–20°C (15°C on average). In the corresponding DNA/DNA duplexes, a single mismatch lowers the T_m by 4–16°C (11°C on average). Because PNA probes can be significantly shorter than DNA probes, their specificity is greater. Fourth, PNA oligomers are resistant to degradation by enzymes, and the lifetime of these compounds is extended both *in vivo* and *in vitro* because nucleases and proteases do not recognize the PNA polyamide backbone with nucleobase sidechains. See, e.g., Ray *et al.*, *FASEB J.* 14(9): 1041-60 (2000); Nielsen *et al.*, *Pharmacol Toxicol.* 86(1): 3-7 (2000); Larsen *et al.*, *Biochim Biophys Acta.* 1489(1): 159-66 (1999); Nielsen, *Curr. Opin. Struct. Biol.* 9(3): 353-7 (1999), and Nielsen, *Curr. Opin. Biotechnol.* 10(1): 71-5 (1999).

Nucleic acid molecules may be modified compared to their native structure throughout the length of the nucleic acid molecule or can be localized to discrete portions thereof. As an example of the latter, chimeric nucleic acids can be synthesized that have discrete DNA and RNA domains and that can be used for targeted gene repair and modified PCR reactions, as further described in, Misra *et al.*, *Biochem.* 37: 1917-1925 (1998); and Finn *et al.*, *Nucl. Acids Res.* 24: 3357-3363 (1996), and U.S. Patent Nos. 5,760,012 and 5,731,181, the disclosures of which are incorporated herein by reference in their entireties.

Unless otherwise specified, nucleic acid molecules of the present invention can include any topological conformation appropriate to the desired use; the term thus explicitly comprehends, among others, single-stranded, double-stranded, triplexed, quadruplexed, partially double-stranded, partially-triplexed, partially-quadruplexed, 5 branched, hairpinned, circular, and padlocked conformations. Padlocked conformations and their utilities are further described in Banér *et al.*, *Curr. Opin. Biotechnol.* 12: 11-15 (2001); Escude *et al.*, *Proc. Natl. Acad. Sci. USA* 14: 96(19):10603-7 (1999); and Nilsson *et al.*, *Science* 265(5181): 2085-8 (1994). Triplexed and quadruplexed conformations, and their utilities, are reviewed in Praseuth *et al.*, *Biochim. Biophys. Acta.* 1489(1): 181-206 10 (1999); Fox, *Curr. Med. Chem.* 7(1): 17-37 (2000); Kochetkova *et al.*, *Methods Mol. Biol.* 130: 189-201 (2000); Chan *et al.*, *J. Mol. Med.* 75(4): 267-82 (1997); Rowley *et al.*, *Mol. Med.* 5(10): 693-700 (1999); Kool, *Annu Rev Biophys Biomol Struct.* 25: 1-28 (1996).

SNP Polymorphisms

Commonly, sequence differences between individuals involve differences in single 15 nucleotide positions. SNPs may account for 90% of human DNA polymorphism. Collins *et al.*, 8 *Genome Res.* 1229-31 (1998). SNPs include single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in a population. In addition, the least frequent allele generally must occur at a frequency of 1% or greater. DNA sequence variants with a reasonably high population frequency are observed 20 approximately every 1,000 nucleotide across the genome, with estimates as high as 1 SNP per 350 base pairs. Wang *et al.*, 280 *Science* 1077-82 (1998); Harding *et al.*, 60 *Am. J. Human Genet.* 772-89 (1997); Taillon-Miller *et al.*, 8 *Genome Res.* 748-54 (1998); Cargill *et al.*, 22 *Nat. Genet.* 231-38 (1999); and Semple *et al.*, 16 *Bioinform. Disc. Note* 735-38 25 (2000). The frequency of SNPs varies with the type and location of the change. In base substitutions, two-thirds of the substitutions involve the C-T and G-A type. This variation in frequency can be related to 5-methylcytosine deamination reactions that occur frequently, particularly at CpG dinucleotides. Regarding location, SNPs occur at a much higher frequency in non-coding regions than in coding regions. Information on over one million variable sequences is already publicly available via the Internet and more such 30 markers are available from commercial providers of genetic information. Kwok and Gu, 5 *Med. Today* 538-53 (1999).

Several definitions of SNPs exist. See, e.g., Brooks, 235 *Gene* 177-86 (1999). As used herein, the term "single nucleotide polymorphism" or "SNP" includes all single base variants, thus including nucleotide insertions and deletions in addition to single nucleotide substitutions. There are two types of nucleotide substitutions. A transition is the
5 replacement of one purine by another purine or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine for a pyrimidine, or vice versa.

Numerous methods exist for detecting SNPs within a nucleotide sequence. A review of many of these methods can be found in Landegren *et al.*, 8 *Genome Res.* 769-76 (1998). For example, a SNP in a genomic sample can be detected by preparing a Reduced
10 Complexity Genome (RCG) from the genomic sample, then analyzing the RCG for the presence or absence of a SNP. See, e.g., WO 00/18960 which is herein incorporated by reference in its entirety. Multiple SNPs in a population of target polynucleotides in parallel can be detected using, for example, the methods of WO 00/50869 which is herein incorporated by reference in its entirety. Other SNP detection methods include the
15 methods of U.S. Pat. Nos. 6,297,018 and 6,322,980 which are herein incorporated by reference in their entirety. Furthermore, SNPs can be detected by restriction fragment length polymorphism (RFLP) analysis. See, e.g., U.S. Pat. Nos. 5,324,631; 5,645,995 which are herein incorporated by reference in their entirety. RFLP analysis of SNPs, however, is limited to cases where the SNP either creates or destroys a restriction enzyme
20 cleavage site. SNPs can also be detected by direct sequencing of the nucleotide sequence of interest. In addition, numerous assays based on hybridization have also been developed to detect SNPs and mismatch distinction by polymerases and ligases. Several web sites provide information about SNPs including Ensembl on the World Wide Web at ensemble.org, Sanger Institute on the World Wide Web at sanger.ac.uk/genetics/exon/,
25 National Center for Biotechnology Information (NCBI) on the World Wide Web at ncbi.nlm.nih.gov/SNP/, The SNP Consortium Ltd. on the World Wide Web at.snp.cshl.org. The chromosomal locations for the compositions disclosed herein are provided below. In addition, one of ordinary skill in the art could use a BLAST against the genome or any of the databases cited above to find the chromosomal location.
30 Another a preferred method to find the genomic coordinates and associated SNPs would be to use the BLAT tool (genome.ucsc.edu, Kent et al. 2001, The Human Genome Browser at UCSC, Genome Research 996-1006 or Kent 2002 BLAT —The BLAST -Like

Alignment Tool Genome Research, 1-9). All web sites above were accessed December 3, 2003.

RNA interference

RNA interference refers to the process of sequence-specific post transcriptional gene silencing in animals mediated by short interfering RNAs (siRNA). Fire *et al.*, 1998, *Nature*, 391, 806. The corresponding process in plants is commonly referred to as post transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post transcriptional gene silencing is thought to be an evolutionarily conserved cellular defense mechanism used to prevent the expression of foreign genes which is commonly shared by diverse flora and phyla. Fire *et al.*, 1999, *Trends Genet.*, 15, 358. Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNA) derived from viral infection or the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA. The presence of dsRNA in cells triggers the RNAi response though a mechanism that has yet to be fully characterized. This mechanism appears to be different from the interferon response that results from dsRNA mediated activation of protein kinase PKR and 2',5'-oligoadenylate synthetase resulting in non-specific cleavage of mRNA by ribonuclease L.

The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNA). Berstein *et al.*, 2001, *Nature*, 409, 363. Short interfering RNAs derived from dicer activity are typically about 21-23 nucleotides in length and comprise about 19 base pair duplexes. Dicer has also been implicated in the excision of 21 and 22 nucleotide small temporal RNAs (stRNA) from precursor RNA of conserved structure that are implicated in translational control.

Hutvagner *et al.*, 2001, *Science*, 293, 834. The RNAi response also features an endonuclease complex containing a siRNA, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex. Elbashir *et al.*, 2001, *Genes Dev.*, 15, 188.

Short interfering RNA mediated RNAi has been studied in a variety of systems. Fire *et al.*, 1998, *Nature*, 391, 806, were the first to observe RNAi in C. Elegans. Wianny and Goetz, 1999, *Nature Cell Biol.*, 2, 70, describe RNAi mediated by dsRNA in mouse embryos. Hammond *et al.*, 2000, *Nature*, 404, 293, describe RNAi in Drosophila cells 5 transfected with dsRNA. Elbashir *et al.*, 2001, *Nature*, 411, 494, describe RNAi induced by introduction of duplexes of synthetic 21-nucleotide RNAs in cultured mammalian cells including human embryonic kidney and HeLa cells. Recent work in Drosophila embryonic lysates (Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877) has revealed certain requirements for 10 siRNA length, structure, chemical composition, and sequence that are essential to mediate efficient RNAi activity. These studies have shown that 21 nucleotide siRNA duplexes are most active when containing two nucleotide 3'-overhangs. Furthermore, complete substitution of one or both siRNA strands with 2'-deoxy (2'-H) or 2'-O-methyl nucleotides abolishes RNAi activity, whereas substitution of the 3'-terminal siRNA overhang 15 nucleotides with deoxy nucleotides (2'-H) was shown to be tolerated. Single mismatch sequences in the center of the siRNA duplex were also shown to abolish RNAi activity. In addition, these studies also indicate that the position of the cleavage site in the target RNA is defined by the 5'-end of the siRNA guide sequence rather than the 3'-end. Elbashir et 20 al., 2001, *EMBO J.*, 20, 6877. Other studies have indicated that a 5'-phosphate on the target-complementary strand of a siRNA duplex is required for siRNA activity and that ATP is utilized to maintain the 5'-phosphate moiety on the siRNA. Nykanen *et al.*, 2001, *Cell*, 107, 309.

Studies have shown that replacing the 3'-overhanging segments of a 21-mer siRNA duplex having 2 nucleotide 3' overhangs with deoxyribonucleotides does not have an adverse effect on RNAi activity. Replacing up to 4 nucleotides on each end of the siRNA 25 with deoxyribonucleotides has been reported to be well tolerated whereas complete substitution with deoxyribonucleotides results in no RNAi activity. Elbashir *et al.*, 2001, *EMBO J.*, 20, 6877. In addition, Elbashir *et al.*, supra, also report that substitution of siRNA with 2'-O-methyl nucleotides completely abolishes RNAi activity. Li *et al.*, WO 00/44914, and Beach *et al.*, WO 01/68836 both suggest that siRNA "may include 30 modifications to either the phosphate-sugar back bone or the nucleoside to include at least one of a nitrogen or sulfur heteroatom", however neither application teaches to what extent these modifications are tolerated in siRNA molecules nor provide any examples of such modified siRNA. Kreutzer and Limmer, Canadian Patent Application No. 2,359,180, also

describe certain chemical modifications for use in dsRNA constructs in order to counteract activation of double-stranded RNA-dependent protein kinase PKR, specifically 2'-amino or 2'-O-methyl nucleotides, and nucleotides containing a 2'-O or 4'-C methylene bridge. However, Kreutzer and Limmer similarly fail to show to what extent these modifications
5 are tolerated in siRNA molecules nor do they provide any examples of such modified siRNA.

Parrish et al., 2000, *Molecular Cell*, 6, 1977-1087, tested certain chemical modifications targeting the unc-22 gene in *C. elegans* using long (>25 nt) siRNA transcripts. The authors describe the introduction of thiophosphate residues into these
10 siRNA transcripts by incorporating thiophosphate nucleotide analogs with T7 and T3 RNA polymerase and observed that "RNAs with two [phosphorothioate] modified bases also had substantial decreases in effectiveness as RNAi triggers; [phosphorothioate] modification of more than two residues greatly destabilized the RNAs in vitro and we were not able to assay interference activities." Parrish et al. at 1081. The authors also
15 tested certain modifications at the 2'-position of the nucleotide sugar in the long siRNA transcripts and observed that substituting deoxynucleotides for ribonucleotides "produced a substantial decrease in interference activity", especially in the case of Uridine to Thymidine and/or Cytidine to deoxy-Cytidine substitutions. Parrish et al. In addition, the authors tested certain base modifications, including substituting 4-thiouracil, 5-
20 bromouracil, 5-iodouracil, 3-(aminoallyl)uracil for uracil, and inosine for guanosine in sense and antisense strands of the siRNA, and found that whereas 4-thiouracil and 5-bromouracil were all well tolerated, inosine "produced a substantial decrease in interference activity" when incorporated in either strand. Incorporation of 5-iodouracil and 3-(aminoallyl)uracil in the antisense strand resulted in substantial decrease in RNAi
25 activity as well.

Beach et al., WO 01/68836, describes specific methods for attenuating gene expression using endogenously derived dsRNA. Tuschl et al., WO 01/75164, describes a *Drosophila* in vitro RNAi system and the use of specific siRNA molecules for certain functional genomic and certain therapeutic applications; although Tuschl, 2001, *Chem.
30 Biochem.*, 2, 239-245, doubts that RNAi can be used to cure genetic diseases or viral infection due "to the danger of activating interferon response". Li et al., WO 00/44914, describes the use of specific dsRNAs for use in attenuating the expression of certain target

genes. Zernicka-Goetz et al., WO 01/36646, describes certain methods for inhibiting the expression of particular genes in mammalian cells using certain dsRNA molecules. Fire et al., WO 99/32619, U.S. Patent No. 6,506,559, the contents of which are hereby incorporated by reference in their entirety, describes particular methods for introducing certain dsRNA molecules into cells for use in inhibiting gene expression. Plaetinck et al., WO 00/01846, describes certain methods for identifying specific genes responsible for conferring a particular phenotype in a cell using specific dsRNA molecules. Mello et al., WO 01/29058, describes the identification of specific genes involved in dsRNA mediated RNAi. Deschamps Depaillette et al., International PCT Publication No. WO 99/07409, describes specific compositions consisting of particular dsRNA molecules combined with certain anti-viral agents. Driscoll et al., International PCT Publication No. WO 01/49844, describes specific DNA constructs for use in facilitating gene silencing in targeted organisms. Parrish et al., 2000, Molecular Cell, 6, 1977-1087, describes specific chemically modified siRNA constructs targeting the unc-22 gene of C. elegans. Tuschl et al., International PCT Publication No. WO 02/44321, describe certain synthetic siRNA constructs.

Methods for Using Nucleic Acid Molecules as Probes and Primers

The isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize, and quantify hybridizing nucleic acids in, and isolate hybridizing nucleic acids from, both genomic and transcript-derived nucleic acid samples. When free in solution, such probes are typically, but not invariably, detectably labeled; bound to a substrate, as in a microarray, such probes are typically, but not invariably unlabeled.

In one embodiment, the isolated nucleic acid molecules of the present invention can be used as probes to detect and characterize gross alterations in the gene of a BSNA, such as deletions, insertions, translocations, and duplications of the BSNA genomic locus through fluorescence *in situ* hybridization (FISH) to chromosome spreads. *See, e.g.*, Andreeff et al. (eds.), Introduction to Fluorescence In Situ Hybridization: Principles and Clinical Applications, John Wiley & Sons (1999). The isolated nucleic acid molecules of the present invention can be used as probes to assess smaller genomic alterations using, e.g., Southern blot detection of restriction fragment length polymorphisms. The isolated nucleic acid molecules of the present invention can be used as probes to isolate genomic

clones that include a nucleic acid molecule of the present invention, which thereafter can be restriction mapped and sequenced to identify deletions, insertions, translocations, and substitutions (single nucleotide polymorphisms, SNPs) at the sequence level.

Alternatively, detection techniques such as molecular beacons may be used, see Kostrikis

5 *et al. Science* 279:1228-1229 (1998).

The isolated nucleic acid molecules of the present invention can also be used as probes to detect, characterize, and quantify BSNA in, and isolate BSNA from, transcript-derived nucleic acid samples. In one embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by length,

10 and quantify mRNA by Northern blot of total or poly-A⁺- selected RNA samples. In another embodiment, the isolated nucleic acid molecules of the present invention can be used as hybridization probes to detect, characterize by location, and quantify mRNA by *in situ* hybridization to tissue sections. See, e.g., Schwarchzacher *et al.*, In Situ Hybridization, Springer-Verlag New York (2000). In another preferred embodiment, the

15 isolated nucleic acid molecules of the present invention can be used as hybridization probes to measure the representation of clones in a cDNA library or to isolate hybridizing nucleic acid molecules acids from cDNA libraries, permitting sequence level characterization of mRNAs that hybridize to BSNAs, including, without limitations, identification of deletions, insertions, substitutions, truncations, alternatively spliced forms

20 and single nucleotide polymorphisms. In yet another preferred embodiment, the nucleic acid molecules of the instant invention may be used in microarrays.

All of the aforementioned probe techniques are well within the skill in the art, and are described at greater length in standard texts such as Sambrook (2001), *supra*; Ausubel (1999), *supra*; and Walker *et al.* (eds.), The Nucleic Acids Protocols Handbook, Humana

25 Press (2000).

In another embodiment, a nucleic acid molecule of the invention may be used as a probe or primer to identify and/or amplify a second nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of the invention. In this embodiment, it is preferred that the probe or primer be derived from a nucleic acid molecule encoding a

30 BSP. More preferably, the probe or primer is derived from a nucleic acid molecule encoding a polypeptide having an amino acid sequence of SEQ ID NO: 96-232. Also preferred are probes or primers derived from a BSNA. More preferred are probes or

primers derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95.

In general, a probe or primer is at least 10 nucleotides in length, more preferably at least 12, more preferably at least 14 and even more preferably at least 16 or 17 nucleotides 5 in length. In an even more preferred embodiment, the probe or primer is at least 18 nucleotides in length, even more preferably at least 20 nucleotides and even more preferably at least 22 nucleotides in length. Primers and probes may also be longer in length. For instance, a probe or primer may be 25 nucleotides in length, or may be 30, 40 or 50 nucleotides in length. Methods of performing nucleic acid hybridization using 10 oligonucleotide probes are well known in the art. See, e.g., Sambrook *et al.*, 1989, *supra*, Chapter 11 and pp. 11.31-11.32 and 11.40-11.44, which describes radiolabeling of short probes, and pp. 11.45-11.53, which describe hybridization conditions for oligonucleotide probes, including specific conditions for probe hybridization (pp. 11.50-11.51).

Methods of performing primer-directed amplification are also well known in the 15 art. Methods for performing the polymerase chain reaction (PCR) are compiled, *inter alia*, in McPherson, PCR Basics: From Background to Bench, Springer Verlag (2000); Innis *et al.* (eds.), PCR Applications: Protocols for Functional Genomics, Academic Press (1999); Gelfand *et al.* (eds.), PCR Strategies, Academic Press (1998); Newton *et al.*, PCR, Springer-Verlag New York (1997); Burke (ed.), PCR: Essential Techniques, John Wiley 20 & Son Ltd (1996); White (ed.), PCR Cloning Protocols: From Molecular Cloning to Genetic Engineering, Vol. 67, Humana Press (1996); and McPherson *et al.* (eds.), PCR 2: A Practical Approach, Oxford University Press, Inc. (1995). Methods for performing RT-PCR are collected, e.g., in Siebert *et al.* (eds.), Gene Cloning and Analysis by RT-PCR, Eaton Publishing Company/Bio Techniques Books Division, 1998; and Siebert (ed.), PCR 25 Technique:RT-PCR, Eaton Publishing Company/ BioTechniques Books (1995).

PCR and hybridization methods may be used to identify and/or isolate nucleic acid molecules of the present invention including allelic variants, homologous nucleic acid molecules and fragments. PCR and hybridization methods may also be used to identify, amplify and/or isolate nucleic acid molecules of the present invention that encode 30 homologous proteins, analogs, fusion proteins or muteins of the invention. Nucleic acid primers as described herein can be used to prime amplification of nucleic acid molecules of the invention, using transcript-derived or genomic DNA as the template.

These nucleic acid primers can also be used, for example, to prime single base extension (SBE) for SNP detection (*See, e.g.*, U.S. Pat. No. 6,004,744, the disclosure of which is incorporated herein by reference in its entirety).

Isothermal amplification approaches, such as rolling circle amplification, are also 5 now well-described. *See, e.g.*, Schweitzer *et al.*, *Curr. Opin. Biotechnol.* 12(1): 21-7 (2001); International Patent publications WO 97/19193 and WO 00/15779, and U.S. Patent Nos. 5,854,033 and 5,714,320, the disclosures of which are incorporated herein by reference in their entireties. Rolling circle amplification can be combined with other techniques to facilitate SNP detection. *See, e.g.*, Lizardi *et al.*, *Nature Genet.* 19(3): 10 225-32 (1998).

Nucleic acid molecules of the present invention may be bound to a substrate either covalently or noncovalently. The substrate can be porous or solid, planar or non-planar, unitary or distributed. The bound nucleic acid molecules may be used as hybridization probes, and may be labeled or unlabeled. In a preferred embodiment, the bound nucleic 15 acid molecules are unlabeled.

In one embodiment, the nucleic acid molecule of the present invention is bound to a porous substrate, *e.g.*, a membrane, typically comprising nitrocellulose, nylon, or positively charged derivatized nylon. The nucleic acid molecule of the present invention can be used to detect a hybridizing nucleic acid molecule that is present within a labeled 20 nucleic acid sample, *e.g.*, a sample of transcript-derived nucleic acids. In another embodiment, the nucleic acid molecule is bound to a solid substrate, including, without limitation, glass, amorphous silicon, crystalline silicon or plastics. Examples of plastics include, without limitation, polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, 25 polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, nitrocellulose, or mixtures thereof. The solid substrate may be any shape, including rectangular, disk-like and spherical. In a preferred embodiment, the solid substrate is a microscope slide or slide-shaped substrate.

The nucleic acid molecule of the present invention can be attached covalently to a 30 surface of the support substrate or applied to a derivatized surface in a chaotropic agent that facilitates denaturation and adherence by presumed noncovalent interactions, or some combination thereof. The nucleic acid molecule of the present invention can be bound to a substrate to which a plurality of other nucleic acids are concurrently bound, hybridization

to each of the plurality of bound nucleic acids being separately detectable. At low density, e.g. on a porous membrane, these substrate-bound collections are typically denominated macroarrays; at higher density, typically on a solid support, such as glass, these substrate bound collections of plural nucleic acids are colloquially termed microarrays. As used
5 herein, the term microarray includes arrays of all densities. It is, therefore, another aspect of the invention to provide microarrays that comprise one or more of the nucleic acid molecules of the present invention.

In yet another embodiment, the invention is directed to single exon probes based on the BSNAAs disclosed herein.

10 *Expression Vectors, Host Cells and Recombinant Methods of Producing Polypeptides*

Another aspect of the present invention provides vectors that comprise one or more of the isolated nucleic acid molecules of the present invention, and host cells in which such vectors have been introduced.

15 The vectors can be used, *inter alia*, for propagating the nucleic acid molecules of the present invention in host cells (cloning vectors), for shuttling the nucleic acid molecules of the present invention between host cells derived from disparate organisms (shuttle vectors), for inserting the nucleic acid molecules of the present invention into host cell chromosomes (insertion vectors), for expressing sense or antisense RNA transcripts of
20 the nucleic acid molecules of the present invention *in vitro* or within a host cell, and for expressing polypeptides encoded by the nucleic acid molecules of the present invention, alone or as fusion proteins with heterologous polypeptides (expression vectors). Vectors are by now well known in the art, and are described, *inter alia*, in Jones *et al.* (eds.), Vectors: Cloning Applications: Essential Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Jones *et al.* (eds.), Vectors: Expression Systems: Essential
25 Techniques (Essential Techniques Series), John Wiley & Son Ltd. (1998); Gacesa *et al.*, Vectors: Essential Data, John Wiley & Sons Ltd. (1995); Cid-Arregui (eds.), Viral
Techniques: Basic Science and Gene Therapy, Eaton Publishing Co. (2000); Sambrook (2001), *supra*; Ausubel (1999), *supra*. Furthermore, a variety of vectors are available
30 commercially. Use of existing vectors and modifications thereof are well within the skill in the art. Thus, only basic features need be described here.

Nucleic acid sequences may be expressed by operatively linking them to an expression control sequence in an appropriate expression vector and employing that expression vector to transform an appropriate unicellular host. Expression control sequences are sequences that control the transcription, post-transcriptional events and 5 translation of nucleic acid sequences. Such operative linking of a nucleic acid sequence of this invention to an expression control sequence, of course, includes, if not already part of the nucleic acid sequence, the provision of a translation initiation codon, ATG or GTG, in the correct reading frame upstream of the nucleic acid sequence.

A wide variety of host/expression vector combinations may be employed in 10 expressing the nucleic acid sequences of this invention. Useful expression vectors, for example, may consist of segments of chromosomal, non-chromosomal and synthetic nucleic acid sequences.

In one embodiment, prokaryotic cells may be used with an appropriate vector. Prokaryotic host cells are often used for cloning and expression. In a preferred 15 embodiment, prokaryotic host cells include *E. coli*, *Pseudomonas*, *Bacillus* and *Streptomyces*. In a preferred embodiment, bacterial host cells are used to express the nucleic acid molecules of the instant invention. Useful expression vectors for bacterial hosts include bacterial plasmids, such as those from *E. coli*, *Bacillus* or *Streptomyces*, including pBluescript, pGEX-2T, pUC vectors, col E1, pCR1, pBR322, pMB9 and their 20 derivatives, wider host range plasmids, such as RP4, phage DNAs, e.g., the numerous derivatives of phage lambda, e.g., NM989, λGT10 and λGT11, and other phages, e.g., M13 and filamentous single-stranded phage DNA. Where *E. coli* is used as host, selectable markers are, analogously, chosen for selectivity in gram negative bacteria: e.g., typical markers confer resistance to antibiotics, such as ampicillin, tetracycline, 25 chloramphenicol, kanamycin, streptomycin and zeocin; auxotrophic markers can also be used.

In other embodiments, eukaryotic host cells, such as yeast, insect, mammalian or plant cells, may be used. Yeast cells, typically *S. cerevisiae*, are useful for eukaryotic 30 genetic studies, due to the ease of targeting genetic changes by homologous recombination and the ability to easily complement genetic defects using recombinantly expressed proteins. Yeast cells are useful for identifying interacting protein components, e.g. through use of a two-hybrid system. In a preferred embodiment, yeast cells are useful for protein expression. Vectors of the present invention for use in yeast will typically, but not

invariably, contain an origin of replication suitable for use in yeast and a selectable marker that is functional in yeast. Yeast vectors include Yeast Integrating plasmids (e.g., YIp5) and Yeast Replicating plasmids (the YRp and YEp series plasmids), Yeast Centromere plasmids (the YCp series plasmids), Yeast Artificial Chromosomes (YACs) which are based on yeast linear plasmids, denoted YLp, pGPD-2, 2 μ plasmids and derivatives thereof, and improved shuttle vectors such as those described in Gietz *et al.*, *Gene*, 74: 5 527-34 (1988) (YIplac, YEplac and YCplac). Selectable markers in yeast vectors include a variety of auxotrophic markers, the most common of which are (in *Saccharomyces cerevisiae*) URA3, HIS3, LEU2, TRP1 and LYS2, which complement specific 10 auxotrophic mutations, such as ura3-52, his3-D1, leu2-D1, trp1-D1 and lys2-201.

Insect cells may be chosen for high efficiency protein expression. Where the host cells are from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSFTM cells (Protein Sciences Corp., Meriden, CT, USA), the vector replicative strategy is typically based upon the baculovirus life cycle. Typically, baculovirus transfer vectors are used to 15 replace the wild-type AcMNPV polyhedrin gene with a heterologous gene of interest. Sequences that flank the polyhedrin gene in the wild-type genome are positioned 5' and 3' of the expression cassette on the transfer vectors. Following co-transfection with AcMNPV DNA, a homologous recombination event occurs between these sequences resulting in a recombinant virus carrying the gene of interest and the polyhedrin or p10 20 promoter. Selection can be based upon visual screening for lacZ fusion activity.

The host cells may also be mammalian cells, which are particularly useful for expression of proteins intended as pharmaceutical agents, and for screening of potential agonists and antagonists of a protein or a physiological pathway. Mammalian vectors intended for autonomous extrachromosomal replication will typically include a viral 25 origin, such as the SV40 origin (for replication in cell lines expressing the large T-antigen, such as COS1 and COS7 cells), the papillomavirus origin, or the EBV origin for long term episomal replication (for use, e.g., in 293-EBNA cells, which constitutively express the EBV EBNA-1 gene product and adenovirus E1A). Vectors intended for integration, and thus replication as part of the mammalian chromosome, can, but need not, include an 30 origin of replication functional in mammalian cells, such as the SV40 origin. Vectors based upon viruses, such as adenovirus, adeno-associated virus, vaccinia virus, and various mammalian retroviruses, will typically replicate according to the viral replicative strategy. Selectable markers for use in mammalian cells include, but are not limited to,

resistance to neomycin (G418), blasticidin, hygromycin and zeocin, and selection based upon the purine salvage pathway using HAT medium.

Expression in mammalian cells can be achieved using a variety of plasmids, including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (*e.g.*, vaccinia virus, 5 adeno virus, and baculovirus), episomal virus vectors (*e.g.*, bovine papillomavirus), and retroviral vectors (*e.g.*, murine retroviruses). Useful vectors for insect cells include baculoviral vectors and pVL 941.

Plant cells can also be used for expression, with the vector replicon typically derived from a plant virus (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, 10 TMV) and selectable markers chosen for suitability in plants.

It is known that codon usage of different host cells may be different. For example, a plant cell and a human cell may exhibit a difference in codon preference for encoding a particular amino acid. As a result, human mRNA may not be efficiently translated in a plant, bacteria or insect host cell. Therefore, another embodiment of this invention is 15 directed to codon optimization. The codons of the nucleic acid molecules of the invention may be modified to resemble, as much as possible, genes naturally contained within the host cell without altering the amino acid sequence encoded by the nucleic acid molecule.

Any of a wide variety of expression control sequences may be used in these vectors to express the nucleic acid molecules of this invention. Such useful expression 20 control sequences include the expression control sequences associated with structural genes of the foregoing expression vectors. Expression control sequences that control transcription include, *e.g.*, promoters, enhancers and transcription termination sites. Expression control sequences in eukaryotic cells that control post-transcriptional events include splice donor and acceptor sites and sequences that modify the half-life of the 25 transcribed RNA, *e.g.*, sequences that direct poly(A) addition or binding sites for RNA-binding proteins. Expression control sequences that control translation include ribosome binding sites, sequences which direct targeted expression of the polypeptide to or within particular cellular compartments, and sequences in the 5' and 3' untranslated regions that modify the rate or efficiency of translation.

30 Examples of useful expression control sequences for a prokaryote, *e.g.*, *E. coli*, will include a promoter, often a phage promoter, such as phage lambda pL promoter, the trc promoter, a hybrid derived from the trp and lac promoters, the bacteriophage T7 promoter (in *E. coli* cells engineered to express the T7 polymerase), the TAC or TRC

system, the major operator and promoter regions of phage lambda, the control regions of fd coat protein, and the araBAD operon. Prokaryotic expression vectors may further include transcription terminators, such as the *aspA* terminator, and elements that facilitate translation, such as a consensus ribosome binding site and translation termination codon,

5 Schomer *et al.*, *Proc. Natl. Acad. Sci. USA* 83: 8506-8510 (1986).

Expression control sequences for yeast cells, typically *S. cerevisiae*, will include a yeast promoter, such as the CYC1 promoter, the GAL1 promoter, the GAL10 promoter, ADH1 promoter, the promoters of the yeast α -mating system, or the GPD promoter, and will typically have elements that facilitate transcription termination, such as the

10 transcription termination signals from the CYC1 or ADH1 gene.

Expression vectors useful for expressing proteins in mammalian cells will include a promoter active in mammalian cells. These promoters include, but are not limited to, those derived from mammalian viruses, such as the enhancer-promoter sequences from the immediate early gene of the human cytomegalovirus (CMV), the enhancer-promoter

15 sequences from the Rous sarcoma virus long terminal repeat (RSV LTR), the enhancer-promoter from SV40 and the early and late promoters of adenovirus. Other expression control sequences include the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase. Other expression control sequences include those from the gene comprising the BSNA of interest. Often, expression is enhanced by
20 incorporation of polyadenylation sites, such as the late SV40 polyadenylation site and the polyadenylation signal and transcription termination sequences from the bovine growth hormone (BGH) gene, and ribosome binding sites. Furthermore, vectors can include introns, such as intron II of rabbit β -globin gene and the SV40 splice elements.

Preferred nucleic acid vectors also include a selectable or amplifiable marker gene
25 and means for amplifying the copy number of the gene of interest. Such marker genes are well known in the art. Nucleic acid vectors may also comprise stabilizing sequences (e.g., ori- or ARS-like sequences and telomere-like sequences), or may alternatively be designed to favor directed or non-directed integration into the host cell genome. In a preferred embodiment, nucleic acid sequences of this invention are inserted in frame into an
30 expression vector that allows a high level expression of an RNA which encodes a protein comprising the encoded nucleic acid sequence of interest. Nucleic acid cloning and sequencing methods are well known to those of skill in the art and are described in an assortment of laboratory manuals, including Sambrook (1989), *supra*, Sambrook (2000),

supra; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Product information from manufacturers of biological, chemical and immunological reagents also provide useful information.

- Expression vectors may be either constitutive or inducible. Inducible vectors
- 5 include either naturally inducible promoters, such as the trc promoter, which is regulated by the lac operon, and the pL promoter, which is regulated by tryptophan, the MMTV-LTR promoter, which is inducible by dexamethasone, or can contain synthetic promoters and/or additional elements that confer inducible control on adjacent promoters. Examples of inducible synthetic promoters are the hybrid Plac/ara-1 promoter and the
- 10 PLtetO-1 promoter. The PLtetO-1 promoter takes advantage of the high expression levels from the PL promoter of phage lambda, but replaces the lambda repressor sites with two copies of operator 2 of the Tn10 tetracycline resistance operon, causing this promoter to be tightly repressed by the Tet repressor protein and induced in response to tetracycline (Tc) and Tc derivatives such as anhydrotetracycline. Vectors may also be inducible
- 15 because they contain hormone response elements, such as the glucocorticoid response element (GRE) and the estrogen response element (ERE), which can confer hormone inducibility where vectors are used for expression in cells having the respective hormone receptors. To reduce background levels of expression, elements responsive to ecdysone, an insect hormone, can be used instead, with coexpression of the ecdysone receptor.

- 20 In one embodiment of the invention, expression vectors can be designed to fuse the expressed polypeptide to small protein tags that facilitate purification and/or visualization. Such tags include a polyhistidine tag that facilitates purification of the fusion protein by immobilized metal affinity chromatography, for example using NiNTA resin (Qiagen Inc., Valencia, CA, USA) or TALON™ resin (cobalt immobilized affinity chromatography
- 25 medium, Clontech Labs, Palo Alto, CA, USA). The fusion protein can include a chitin-binding tag and self-excising intein, permitting chitin-based purification with self-removal of the fused tag (IMPACT™ system, New England Biolabs, Inc., Beverley, MA, USA). Alternatively, the fusion protein can include a calmodulin-binding peptide tag, permitting purification by calmodulin affinity resin (Stratagene, La Jolla, CA, USA), or a specifically
- 30 excisable fragment of the biotin carboxylase carrier protein, permitting purification of *in vivo* biotinylated protein using an avidin resin and subsequent tag removal (Promega, Madison, WI, USA). As another useful alternative, the polypeptides of the present invention can be expressed as a fusion to glutathione-S-transferase, the affinity and

specificity of binding to glutathione permitting purification using glutathione affinity resins, such as Glutathione-Superflow Resin (Clontech Laboratories, Palo Alto, CA, USA), with subsequent elution with free glutathione. Other tags include, for example, the Xpress epitope, detectable by anti-Xpress antibody (Invitrogen, Carlsbad, CA, USA), a 5 myc tag, detectable by anti-myc tag antibody, the V5 epitope, detectable by anti-V5 antibody (Invitrogen, Carlsbad, CA, USA), FLAG® epitope, detectable by anti-FLAG® antibody (Stratagene, La Jolla, CA, USA), and the HA epitope, detectable by anti-HA antibody.

For secretion of expressed polypeptides, vectors can include appropriate sequences 10 that encode secretion signals, such as leader peptides. For example, the pSecTag2 vectors (Invitrogen, Carlsbad, CA, USA) are 5.2 kb mammalian expression vectors that carry the secretion signal from the V-J2-C region of the mouse Ig kappa-chain for efficient secretion of recombinant proteins from a variety of mammalian cell lines.

Expression vectors can also be designed to fuse proteins encoded by the 15 heterologous nucleic acid insert to polypeptides that are larger than purification and/or identification tags. Useful protein fusions include those that permit display of the encoded protein on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, such as those that have a green fluorescent protein (GFP)-like chromophore, fusions to the IgG Fc region, and fusions for use in two hybrid systems.

20 Vectors for phage display fuse the encoded polypeptide to, e.g., the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13. See Barbas *et al.*, Phage Display: A Laboratory Manual, Cold Spring Harbor Laboratory Press (2001); Kay *et al.* (eds.), Phage Display of Peptides and Proteins: A Laboratory Manual, Academic Press, Inc., (1996); Abelson *et al.* (eds.), Combinatorial 25 Chemistry (Methods in Enzymology, Vol. 267) Academic Press (1996). Vectors for yeast display, e.g. the pYD1 yeast display vector (Invitrogen, Carlsbad, CA, USA), use the α-agglutinin yeast adhesion receptor to display recombinant protein on the surface of *S. cerevisiae*. Vectors for mammalian display, e.g., the pDisplay™ vector (Invitrogen, Carlsbad, CA, USA), target recombinant proteins using an N-terminal cell surface 30 targeting signal and a C-terminal transmembrane anchoring domain of platelet derived growth factor receptor.

A wide variety of vectors now exist that fuse proteins encoded by heterologous nucleic acids to the chromophore of the substrate-independent, intrinsically fluorescent

green fluorescent protein from *Aequorea victoria* (“GFP”) and its variants. The GFP-like chromophore can be selected from GFP-like chromophores found in naturally occurring proteins, such as *A. victoria* GFP (GenBank accession number AAA27721), *Renilla reniformis* GFP, FP583 (GenBank accession no. AF168419) (DsRed), FP593 (AF272711),
5 FP483 (AF168420), FP484 (AF168424), FP595 (AF246709), FP486 (AF168421), FP538 (AF168423), and FP506 (AF168422), and need include only so much of the native protein as is needed to retain the chromophore’s intrinsic fluorescence. Methods for determining the minimal domain required for fluorescence are known in the art. See Li *et al.*, *J. Biol. Chem.* 272: 28545-28549 (1997). Alternatively, the GFP-like chromophore can be
10 selected from GFP-like chromophores modified from those found in nature. The methods for engineering such modified GFP-like chromophores and testing them for fluorescence activity, both alone and as part of protein fusions, are well known in the art. See Heim *et al.*, *Curr. Biol.* 6: 178-182 (1996) and Palm *et al.*, *Methods Enzymol.* 302: 378-394 (1999). A variety of such modified chromophores are now commercially available and can readily
15 be used in the fusion proteins of the present invention. These include EGFP (“enhanced GFP”), EBFP (“enhanced blue fluorescent protein”), BFP2, EYFP (“enhanced yellow fluorescent protein”), ECFP (“enhanced cyan fluorescent protein”) or Citrine. EGFP (see, e.g., Cormack *et al.*, *Gene* 173: 33-38 (1996); U.S. Patent Nos. 6,090,919 and 5,804,387, the disclosures of which are incorporated herein by reference in their entireties) is found
20 on a variety of vectors, both plasmid and viral, which are available commercially (Clontech Labs, Palo Alto, CA, USA); EBFP is optimized for expression in mammalian cells whereas BFP2, which retains the original jellyfish codons, can be expressed in bacteria (see, e.g., Heim *et al.*, *Curr. Biol.* 6: 178-182 (1996) and Cormack *et al.*, *Gene* 173: 33-38 (1996)). Vectors containing these blue-shifted variants are available from
25 Clontech Labs (Palo Alto, CA, USA). Vectors containing EYFP, ECFP (see, e.g., Heim *et al.*, *Curr. Biol.* 6: 178-182 (1996); Miyawaki *et al.*, *Nature* 388: 882-887 (1997)) and Citrine (see, e.g., Heikal *et al.*, *Proc. Natl. Acad. Sci. USA* 97: 11996-12001 (2000)) are also available from Clontech Labs. The GFP-like chromophore can also be drawn from other modified GFPs, including those described in U.S. Patent Nos. 6,124,128; 6,096,865;
30 6,090,919; 6,066,476; 6,054,321; 6,027,881; 5,968,750; 5,874,304; 5,804,387; 5,777,079; 5,741,668; and 5,625,048, the disclosures of which are incorporated herein by reference in their entireties. See also Conn (ed.), Green Fluorescent Protein (Methods in Enzymology, Vol. 302), Academic Press, Inc. (1999); Yang, *et al.*, *J Biol Chem.* 273: 8212-6 (1998);

Bevis *et al.*, *Nature Biotechnology*, 20:83-7 (2002). The GFP-like chromophore of each of these GFP variants can usefully be included in the fusion proteins of the present invention.

Fusions to the IgG Fc region increase serum half-life of protein pharmaceutical products through interaction with the FcRn receptor (also denominated the FcRp receptor and the Brambell receptor, FcRb), further described in International Patent Application Nos. WO 97/43316, WO 97/34631, WO 96/32478, and WO 96/18412, the disclosures of which are incorporated herein by reference in their entireties.

For long-term, high-yield recombinant production of the polypeptides of the present invention, stable expression is preferred. Stable expression is readily achieved by integration into the host cell genome of vectors having selectable markers, followed by selection of these integrants. Vectors such as pUB6/V5-His A, B, and C (Invitrogen, Carlsbad, CA, USA) are designed for high-level stable expression of heterologous proteins in a wide range of mammalian tissue types and cell lines. pUB6/V5-His uses the promoter/enhancer sequence from the human ubiquitin C gene to drive expression of recombinant proteins: expression levels in 293, CHO, and NIH3T3 cells are comparable to levels from the CMV and human EF-1a promoters. The bsd gene permits rapid selection of stably transfected mammalian cells with the potent antibiotic blasticidin.

Replication incompetent retroviral vectors, typically derived from Moloney murine leukemia virus, also are useful for creating stable transfectants having integrated provirus. The highly efficient transduction machinery of retroviruses, coupled with the availability of a variety of packaging cell lines such as RetroPack™ PT 67, EcoPack2™-293, AmphiPack-293, and GP2-293 cell lines (all available from Clontech Laboratories, Palo Alto, CA, USA) allow a wide host range to be infected with high efficiency; varying the multiplicity of infection readily adjusts the copy number of the integrated provirus.

Of course, not all vectors and expression control sequences will function equally well to express the nucleic acid molecules of this invention. Neither will all hosts function equally well with the same expression system. However, one of skill in the art may make a selection among these vectors, expression control sequences and hosts without undue experimentation and without departing from the scope of this invention. For example, in selecting a vector, the host must be considered because the vector must be replicated in it. The vector's copy number, the ability to control that copy number, the ability to control integration, if any, and the expression of any other proteins encoded by the vector, such as

an antibiotic or other selection marker, should also be considered. The present invention further includes host cells comprising the vectors of the present invention, either present episomally within the cell or integrated, in whole or in part, into the host cell chromosome.

Among other considerations, some of which are described above, a host cell strain may be

- 5 chosen for its ability to process the expressed polypeptide in the desired fashion. Such post-translational modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation, and it is an aspect of the present invention to provide BSPs with such post-translational modifications.

10 In selecting an expression control sequence, a variety of factors should also be considered. These include, for example, the relative strength of the sequence, its controllability, and its compatibility with the nucleic acid molecules of this invention, particularly with regard to potential secondary structures. Unicellular hosts should be selected by consideration of their compatibility with the chosen vector, the toxicity of the

15 product coded for by the nucleic acid sequences of this invention, their secretion characteristics, their ability to fold the polypeptide correctly, their fermentation or culture requirements, and the ease of purification from them of the products coded for by the nucleic acid molecules of this invention.

The recombinant nucleic acid molecules and more particularly, the expression vectors of this invention may be used to express the polypeptides of this invention as recombinant polypeptides in a heterologous host cell. The polypeptides of this invention may be full-length or less than full-length polypeptide fragments recombinantly expressed from the nucleic acid molecules according to this invention. Such polypeptides include analogs, derivatives and muteins that may or may not have biological activity.

25 Vectors of the present invention will also often include elements that permit *in vitro* transcription of RNA from the inserted heterologous nucleic acid. Such vectors typically include a phage promoter, such as that from T7, T3, or SP6, flanking the nucleic acid insert. Often two different such promoters flank the inserted nucleic acid, permitting separate *in vitro* production of both sense and antisense strands.

30 Transformation and other methods of introducing nucleic acids into a host cell (e.g., conjugation, protoplast transformation or fusion, transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion) can be accomplished by a variety of methods which are

well known in the art (See, for instance, Ausubel, *supra*, and Sambrook *et al.*, *supra*).

Bacterial, yeast, plant or mammalian cells are transformed or transfected with an expression vector, such as a plasmid, a cosmid, or the like, wherein the expression vector comprises the nucleic acid of interest. Alternatively, the cells may be infected by a viral expression vector comprising the nucleic acid of interest. Depending upon the host cell, vector, and method of transformation used, transient or stable expression of the polypeptide will be constitutive or inducible. One having ordinary skill in the art will be able to decide whether to express a polypeptide transiently or stably, and whether to express the protein constitutively or inducibly.

- 10 A wide variety of unicellular host cells are useful in expressing the DNA sequences of this invention. These hosts may include well known eukaryotic and prokaryotic hosts, such as strains of, fungi, yeast, insect cells such as *Spodoptera frugiperda* (SF9), animal cells such as CHO, as well as plant cells in tissue culture. Representative examples of appropriate host cells include, but are not limited to, bacterial
- 15 cells, such as *E. coli*, *Caulobacter crescentus*, *Streptomyces* species, and *Salmonella typhimurium*; yeast cells, such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Pichia methanolica*; insect cell lines, such as those from *Spodoptera frugiperda*, e.g., Sf9 and Sf21 cell lines, and expresSFTM cells (Protein Sciences Corp., Meriden, CT, USA), *Drosophila* S2 cells, and *Trichoplusia ni* High Five® Cells
- 20 (Invitrogen, Carlsbad, CA, USA); and mammalian cells. Typical mammalian cells include BHK cells, BSC 1 cells, BSC 40 cells, BMT 10 cells, VERO cells, COS1 cells, COS7 cells, Chinese hamster ovary (CHO) cells, 3T3 cells, NIH 3T3 cells, 293 cells, HEPG2 cells, HeLa cells, L cells, MDCK cells, HEK293 cells, WI38 cells, murine ES cell lines (e.g., from strains 129/SV, C57/BL6, DBA-1, 129/SVJ), K562 cells, Jurkat cells, and
- 25 BW5147 cells. Other mammalian cell lines are well known and readily available from the American Type Culture Collection (ATCC) (Manassas, VA, USA) and the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Cell Repositories (Camden, NJ, USA). Cells or cell lines derived from breast are particularly preferred because they may provide a more native post-translational
- 30 processing. Particularly preferred are human breast cells.

Particular details of the transfection, expression and purification of recombinant proteins are well documented and are understood by those of skill in the art. Further details on the various technical aspects of each of the steps used in recombinant

production of foreign genes in bacterial cell expression systems can be found in a number of texts and laboratory manuals in the art. See, e.g., Ausubel (1992), *supra*, Ausubel (1999), *supra*, Sambrook (1989), *supra*, and Sambrook (2001), *supra*.

- Methods for introducing the vectors and nucleic acid molecules of the present
5 invention into the host cells are well known in the art; the choice of technique will depend primarily upon the specific vector to be introduced and the host cell chosen.

Nucleic acid molecules and vectors may be introduced into prokaryotes, such as *E. coli*, in a number of ways. For instance, phage lambda vectors will typically be packaged using a packaging extract (e.g., Gigapack® packaging extract, Stratagene, La Jolla, CA, 10 USA), and the packaged virus used to infect *E. coli*.

Plasmid vectors will typically be introduced into chemically competent or electrocompetent bacterial cells. *E. coli* cells can be rendered chemically competent by treatment, e.g., with CaCl₂, or a solution of Mg²⁺, Mn²⁺, Ca²⁺, Rb⁺ or K⁺, dimethyl sulfoxide, dithiothreitol, and hexamine cobalt (III), Hanahan, *J. Mol. Biol.* 166(4):557-80 (1983), and vectors introduced by heat shock. A wide variety of chemically competent strains are also available commercially (e.g., Epicurian Coli® XL10-Gold® Ultracompetent Cells (Stratagene, La Jolla, CA, USA); DH5 α competent cells (Clontech Laboratories, Palo Alto, CA, USA); and TOP10 Chemically Competent *E. coli* Kit (Invitrogen, Carlsbad, CA, USA)). Bacterial cells can be rendered electrocompetent to 20 take up exogenous DNA by electroporation by various pre-pulse treatments; vectors are introduced by electroporation followed by subsequent outgrowth in selected media. An extensive series of protocols is provided by BioRad (Richmond, CA, USA).

Vectors can be introduced into yeast cells by spheroplasting, treatment with lithium salts, electroporation, or protoplast fusion. Spheroplasts are prepared by the action 25 of hydrolytic enzymes such as a snail-gut extract, usually denoted Glusulase or Zymolyase, or an enzyme from *Arthrobacter luteus* to remove portions of the cell wall in the presence of osmotic stabilizers, typically 1 M sorbitol. DNA is added to the spheroplasts, and the mixture is co-precipitated with a solution of polyethylene glycol (PEG) and Ca²⁺. Subsequently, the cells are resuspended in a solution of sorbitol, mixed 30 with molten agar and then layered on the surface of a selective plate containing sorbitol.

For lithium-mediated transformation, yeast cells are treated with lithium acetate to permeabilize the cell wall, DNA is added and the cells are co-precipitated with PEG. The cells are exposed to a brief heat shock, washed free of PEG and lithium acetate, and

subsequently spread on plates containing ordinary selective medium. Increased frequencies of transformation are obtained by using specially-prepared single-stranded carrier DNA and certain organic solvents. Schiestl *et al.*, *Curr. Genet.* 16(5-6): 339-46 (1989).

- 5 For electroporation, freshly-grown yeast cultures are typically washed, suspended in an osmotic protectant, such as sorbitol, mixed with DNA, and the cell suspension pulsed in an electroporation device. Subsequently, the cells are spread on the surface of plates containing selective media. Becker *et al.*, *Methods Enzymol.* 194: 182-187 (1991). The efficiency of transformation by electroporation can be increased over 100-fold by
10 using PEG, single-stranded carrier DNA and cells that are in late log-phase of growth. Larger constructs, such as YACs, can be introduced by protoplast fusion.

Mammalian and insect cells can be directly infected by packaged viral vectors, or transfected by chemical or electrical means. For chemical transfection, DNA can be coprecipitated with CaPO₄ or introduced using liposomal and nonliposomal lipid-based
15 agents. Commercial kits are available for CaPO₄ transfection (CalPhos™ Mammalian Transfection Kit, Clontech Laboratories, Palo Alto, CA, USA), and lipid-mediated transfection can be practiced using commercial reagents, such as LIPOFECTAMINE™ 2000, LIPOFECTAMINE™ Reagent, CELLFECTIN® Reagent, and LIPOFECTIN® Reagent (Invitrogen, Carlsbad, CA, USA), DOTAP Liposomal Transfection Reagent,
20 FuGENE 6, X-tremeGENE Q2, DOSPER, (Roche Molecular Biochemicals, Indianapolis, IN USA), Effectene™, PolyFect®, Superfect® (Qiagen, Inc., Valencia, CA, USA). Protocols for electroporating mammalian cells can be found in, for example, ; Norton *et al.* (eds.), *Gene Transfer Methods: Introducing DNA into Living Cells and Organisms*, BioTechniques Books, Eaton Publishing Co. (2000). Other transfection techniques
25 include transfection by particle bombardment and microinjection. See, e.g., Cheng *et al.*, *Proc. Natl. Acad. Sci. USA* 90(10): 4455-9 (1993); Yang *et al.*, *Proc. Natl. Acad. Sci. USA* 87(24): 9568-72 (1990).

Production of the recombinantly produced proteins of the present invention can optionally be followed by purification.

- 30 Purification of recombinantly expressed proteins is now well within the skill in the art and thus need not be detailed here. See, e.g., Thorner *et al.* (eds.), *Applications of Chimeric Genes and Hybrid Proteins. Part A: Gene Expression and Protein Purification* (Methods in Enzymology, Vol. 326), Academic Press (2000); Harbin (ed.), *Cloning, Gene*

Expression and Protein Purification : Experimental Procedures and Process Rationale, Oxford Univ. Press (2001); Marshak *et al.*, Strategies for Protein Purification and Characterization: A Laboratory Course Manual, Cold Spring Harbor Laboratory Press (1996); and Roe (ed.), Protein Purification Applications, Oxford University Press (2001).

5 Briefly, however, if purification tags have been fused through use of an expression vector that appends such tags, purification can be effected, at least in part, by means appropriate to the tag, such as use of immobilized metal affinity chromatography for polyhistidine tags. Other techniques common in the art include ammonium sulfate fractionation, immunoprecipitation, fast protein liquid chromatography (FPLC), high performance liquid chromatography (HPLC), and preparative gel electrophoresis.

10

Polypeptides, including Fragments Muteins, Homologous Proteins, Allelic Variants, Analogs and Derivatives

Another aspect of the invention relates to polypeptides encoded by the nucleic acid molecules described herein. In a preferred embodiment, the polypeptide is a breast specific polypeptide (BSP). In an even more preferred embodiment, the polypeptide comprises an amino acid sequence of SEQ ID NO:96-232 or is derived from a polypeptide having the amino acid sequence of SEQ ID NO: 96-232. A polypeptide as defined herein may be produced recombinantly, as discussed *supra*, may be isolated from a cell that naturally expresses the protein, or may be chemically synthesized following the teachings 20 of the specification and using methods well known to those having ordinary skill in the art.

Polypeptides of the present invention may also comprise a part or fragment of a BSP. In a preferred embodiment, the fragment is derived from a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 96-232.

Polypeptides of the present invention comprising a part or fragment of an entire BSP may 25 or may not be BSPs. For example, a full-length polypeptide may be breast-specific, while a fragment thereof may be found in other tissues as well as in breast. A polypeptide that is not a BSP, whether it is a fragment, analog, mutein, homologous protein or derivative, is nevertheless useful, especially for immunizing animals to prepare anti-BSP antibodies. In a preferred embodiment, the part or fragment is a BSP. Methods of determining whether a 30 polypeptide of the present invention is a BSP are described *infra*.

Polypeptides of the present invention comprising fragments of at least 6 contiguous amino acids are also useful in mapping B cell and T cell epitopes of the

reference protein. *See, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81: 3998-4002 (1984) and U.S. Patent Nos. 4,708,871 and 5,595,915, the disclosures of which are incorporated herein by reference in their entireties. Because the fragment need not itself be immunogenic, part of an immunodominant epitope, nor even recognized by native antibody, to be useful in such epitope mapping, all fragments of at least 6 amino acids of a polypeptide of the present invention have utility in such a study.

Polypeptides of the present invention comprising fragments of at least 8 contiguous amino acids, often at least 15 contiguous amino acids, are useful as immunogens for raising antibodies that recognize polypeptides of the present invention.

10 *See, e.g.*, Lerner, *Nature* 299: 592-596 (1982); Shinnick *et al.*, *Annu. Rev. Microbiol.* 37: 425-46 (1983); Sutcliffe *et al.*, *Science* 219: 660-6 (1983). As further described in the above-cited references, virtually all 8-mers, conjugated to a carrier, such as a protein, prove immunogenic and are capable of eliciting antibody for the conjugated peptide; accordingly, all fragments of at least 8 amino acids of the polypeptides of the present 15 invention have utility as immunogens.

Polypeptides comprising fragments of at least 8, 9, 10 or 12 contiguous amino acids are also useful as competitive inhibitors of binding of the entire polypeptide, or a portion thereof, to antibodies (as in epitope mapping), and to natural binding partners, such as subunits in a multimeric complex or to receptors or ligands of the subject protein; 20 this competitive inhibition permits identification and separation of molecules that bind specifically to the polypeptide of interest. See U.S. Patent Nos. 5,539,084 and 5,783,674, incorporated herein by reference in their entireties.

The polypeptide of the present invention thus preferably is at least 6 amino acids in length, typically at least 8, 9, 10 or 12 amino acids in length, and often at least 15 amino 25 acids in length. Often, the polypeptide of the present invention is at least 20 amino acids in length, even 25 amino acids, 30 amino acids, 35 amino acids, or 50 amino acids or more in length. Of course, larger polypeptides having at least 75 amino acids, 100 amino acids, or even 150 amino acids are also useful, and at times preferred.

One having ordinary skill in the art can produce fragments by truncating the 30 nucleic acid molecule, *e.g.*, a BSNA, encoding the polypeptide and then expressing it recombinantly. Alternatively, one can produce a fragment by chemically synthesizing a portion of the full-length polypeptide. One may also produce a fragment by enzymatically cleaving either a recombinant polypeptide or an isolated naturally occurring polypeptide.

Methods of producing polypeptide fragments are well known in the art. *See, e.g.*, Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. In one embodiment, a polypeptide comprising only a fragment, preferably a fragment of a BSP, may be produced by chemical or enzymatic cleavage of a BSP

5 polypeptide. In a preferred embodiment, a polypeptide fragment is produced by expressing a nucleic acid molecule of the present invention encoding a fragment, preferably of a BSP, in a host cell.

Polypeptides of the present invention are also inclusive of mutants, fusion proteins,

homologous proteins and allelic variants.

10 A mutant protein, or mutein, may have the same or different properties compared to a naturally occurring polypeptide and comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of a native polypeptide. Small deletions and insertions can often be found that do not alter the function of a protein. Muteins may or may not be breast-specific. Preferably, the

15 mutein is breast-specific. More preferably the mutein is a polypeptide that comprises at least one amino acid insertion, duplication, deletion, rearrangement or substitution compared to the amino acid sequence of SEQ ID NO: 96-232. Accordingly, in a preferred embodiment, the mutein is one that exhibits at least 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more

20 preferably at least 80% sequence identity to a BSP comprising an amino acid sequence of SEQ ID NO: 96-232. In a yet more preferred embodiment, the mutein exhibits at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97%, 98%, 99% or 99.5% sequence identity to a BSP comprising an amino acid sequence of SEQ ID NO: 96-232.

25 A mutein may be produced by isolation from a naturally occurring mutant cell, tissue or organism. A mutein may be produced by isolation from a cell, tissue or organism that has been experimentally mutagenized. Alternatively, a mutein may be produced by chemical manipulation of a polypeptide, such as by altering the amino acid residue to another amino acid residue using synthetic or semi-synthetic chemical techniques. In a

30 preferred embodiment, a mutein is produced from a host cell comprising a mutated nucleic acid molecule compared to the naturally occurring nucleic acid molecule. For instance, one may produce a mutein of a polypeptide by introducing one or more mutations into a nucleic acid molecule of the invention and then expressing it recombinantly. These

mutations may be targeted, in which particular encoded amino acids are altered, or may be untargeted, in which random encoded amino acids within the polypeptide are altered. Muteins with random amino acid alterations can be screened for a particular biological activity or property, particularly whether the polypeptide is breast-specific, as described 5 below. Multiple random mutations can be introduced into the gene by methods well known to the art, e.g., by error-prone PCR, shuffling, oligonucleotide-directed mutagenesis, assembly PCR, sexual PCR mutagenesis, *in vivo* mutagenesis, cassette mutagenesis, recursive ensemble mutagenesis, exponential ensemble mutagenesis and site-specific mutagenesis. Methods of producing muteins with targeted or random amino acid 10 alterations are well known in the art. See, e.g., Sambrook (1989), *supra*; Sambrook (2001), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), as well as U.S. Patent No. 5,223,408, which is herein incorporated by reference in its entirety.

The invention also contemplates polypeptides that are homologous to a polypeptide of the invention. In a preferred embodiment, the polypeptide is homologous 15 to a BSP. In an even more preferred embodiment, the polypeptide is homologous to a BSP selected from the group having an amino acid sequence of SEQ ID NO: 96-232. By homologous polypeptide it is meant one that exhibits significant sequence identity to a BSP, preferably a BSP having an amino acid sequence of SEQ ID NO: 96-232. By significant sequence identity it is meant that the homologous polypeptide exhibits at least 20 50% sequence identity, more preferably at least 60% sequence identity, even more preferably at least 70%, yet more preferably at least 80% sequence identity to a BSP comprising an amino acid sequence of SEQ ID NO: 96-232. More preferred are homologous polypeptides exhibiting at least 85%, more preferably 90%, even more preferably 95% or 96%, and yet more preferably at least 97% or 98% sequence identity to 25 a BSP comprising an amino acid sequence of SEQ ID NO: 96-232. Most preferably, the homologous polypeptide exhibits at least 99%, more preferably 99.5%, even more preferably 99.6%, 99.7%, 99.8% or 99.9% sequence identity to a BSP comprising an amino acid sequence of SEQ ID NO: 96-232. In a preferred embodiment, the amino acid substitutions of the homologous polypeptide are conservative amino acid substitutions as 30 discussed *supra*.

Homologous polypeptides of the present invention also comprise polypeptides encoded by a nucleic acid molecule that selectively hybridizes to a BSNA or an antisense sequence thereof. In this embodiment, it is preferred that the homologous polypeptide be

- encoded by a nucleic acid molecule that hybridizes to a BSNA under low stringency, moderate stringency or high stringency conditions, as defined herein. More preferred is a homologous polypeptide encoded by a nucleic acid sequence which hybridizes to a BSNA selected from the group consisting of SEQ ID NO: 1-95 or a homologous polypeptide 5 encoded by a nucleic acid molecule that hybridizes to a nucleic acid molecule that encodes a BSP, preferably a BSP of SEQ ID NO:96-232 under low stringency, moderate stringency or high stringency conditions, as defined herein.

Homologous polypeptides of the present invention may be naturally occurring and derived from another species, especially one derived from another primate, such as 10 chimpanzee, gorilla, rhesus macaque, or baboon, wherein the homologous polypeptide comprises an amino acid sequence that exhibits significant sequence identity to that of SEQ ID NO: 96-232. The homologous polypeptide may also be a naturally occurring polypeptide from a human, when the BSP is a member of a family of polypeptides. The homologous polypeptide may also be a naturally occurring polypeptide derived from a 15 non-primate, mammalian species, including without limitation, domesticated species, e.g., dog, cat, mouse, rat, rabbit, guinea pig, hamster, cow, horse, goat or pig. The homologous polypeptide may also be a naturally occurring polypeptide derived from a non-mammalian species, such as birds or reptiles. The naturally occurring homologous protein may be isolated directly from humans or other species. Alternatively, the nucleic acid molecule 20 encoding the naturally occurring homologous polypeptide may be isolated and used to express the homologous polypeptide recombinantly. The homologous polypeptide may also be one that is experimentally produced by random mutation of a nucleic acid molecule and subsequent expression of the nucleic acid molecule. Alternatively, the homologous polypeptide may be one that is experimentally produced by directed mutation 25 of one or more codons to alter the encoded amino acid of a BSP. In a preferred embodiment, the homologous polypeptide encodes a polypeptide that is a BSP.

Relatedness of proteins can also be characterized using a second functional test, such as the ability of a first protein competitively to inhibit the binding of a second protein to an antibody. It is, therefore, another aspect of the present invention to provide isolated 30 polypeptides not only identical in sequence to those described with particularity herein, but also to provide isolated polypeptides (“cross-reactive proteins”) that competitively inhibit the binding of antibodies to all or to a portion of the isolated polypeptides of the

present invention. Such competitive inhibition can readily be determined using immunoassays well known in the art.

As discussed above, single nucleotide polymorphisms (SNPs) occur frequently in eukaryotic genomes, and the sequence determined from one individual of a species may 5 differ from other allelic forms present within the population. Thus, polypeptides of the present invention are also inclusive of those encoded by an allelic variant of a nucleic acid molecule encoding a BSP. In this embodiment, it is preferred that the polypeptide be encoded by an allelic variant of a gene that encodes a polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO: 96-232. More preferred is 10 that the polypeptide be encoded by an allelic variant of a gene that has the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1-95.

Polypeptides of the present invention are also inclusive of derivative polypeptides encoded by a nucleic acid molecule according to the instant invention. In this embodiment, it is preferred that the polypeptide be a BSP. Also preferred are derivative 15 polypeptides having an amino acid sequence selected from the group consisting of SEQ ID NO: 96-232 and which has been acetylated, carboxylated, phosphorylated, glycosylated, ubiquitinated or post-translationally modified in another manner. In another preferred embodiment, the derivative has been labeled with, *e.g.*, radioactive isotopes such as ^{125}I , ^{32}P , ^{35}S , and ^3H . In another preferred embodiment, the derivative has been labeled 20 with fluorophores, chemiluminescent agents, enzymes, and antiligands that can serve as specific binding pair members for a labeled ligand.

Polypeptide modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of 25 glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance Creighton, Protein Structure and Molecular Properties, 2nd ed., W. H. Freeman and Company (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, in Johnson (ed.), Posttranslational Covalent Modification of Proteins, pgs. 1-12, Academic Press (1983); 30 Seifter *et al.*, *Meth. Enzymol.* 182: 626-646 (1990) and Rattan *et al.*, *Ann. N.Y. Acad. Sci.* 663: 48-62 (1992).

One may determine whether a polypeptide of the invention is likely to be post-translationally modified by analyzing the sequence of the polypeptide to determine if there

are peptide motifs indicative of sites for post-translational modification. There are a number of computer programs that permit prediction of post-translational modifications. See, e.g., expasy.org (accessed November 11, 2002) of the world wide web, which includes PSORT, for prediction of protein sorting signals and localization sites, SignalP, 5 for prediction of signal peptide cleavage sites, MITOPROT and Predotar, for prediction of mitochondrial targeting sequences, NetOGlyc, for prediction of type O-glycosylation sites in mammalian proteins, big-PI Predictor and DGPI, for prediction of prenylation-anchor and cleavage sites, and NetPhos, for prediction of Ser, Thr and Tyr phosphorylation sites in eukaryotic proteins. Other computer programs, such as those included in GCG, also 10 may be used to determine post-translational modification peptide motifs.

General examples of types of post-translational modifications include, but are not limited to: (Z)-dehydrobutyryne; 1-chondroitin sulfate-L-aspartic acid ester; 1'-glycosyl-L-tryptophan; 1'-phospho-L-histidine; 1-thioglycine; 2'-(S-L-cysteinyl)-L-histidine; 2'-[3-carboxamido (trimethylammonio)propyl]-L-histidine; 2'-alpha-mannosyl-L-tryptophan; 2-methyl-L-glutamine; 2-oxobutanoic acid; 2-pyrrolidone carboxylic acid; 3'-(1'L-histidyl)-L-tyrosine; 3'-(8alpha-FAD)-L-histidine; 3'-(S-L-cysteinyl)-L-tyrosine; 3', 3", 5'-triiodo-L-thyronine; 3'-4'-phospho-L-tyrosine; 3-hydroxy-L-proline; 3'-methyl-L-histidine; 3-methyl-L-lanthionine; 3'-phospho-L-histidine; 4'-(L-tryptophan)-L-tryptophyl quinone; 42 N-cysteinyl-glycosylphosphatidylinositolethanolamine; 43 -(T-L-histidyl)-L-tyrosine; 4-hydroxy-L-arginine; 4-hydroxy-L-lysine; 4-hydroxy-L-proline; 5'-(N6-L-lysine)-L-topoquinone; 5-hydroxy-L-lysine; 5-methyl-L-arginine; alpha-L-microglobulin-Ig alpha complex chromophore; bis-L-cysteinyl bis-L-histidino diiron disulfide; bis-L--cysteinyl-L-N3'-histidino-L-serinyI tetrairon' tetrasulfide; chondroitin sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-alanine; D-allo-isoleucine; D-asparagine; 25 dehydroalanine; dehydrotyrosine; dermatan 4-sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; D-glucuronyl-N-glycine; dipyrrolylmethanemethyl-L-cysteine; D-leucine; D-methionine; D-phenylalanine; D-serine; D-tryptophan; glycine amide; glycine oxazolecarboxylic acid; glycine thiazolecarboxylic acid; heme P450-bis-L-cysteine-L-tyrosine; heme-bis-L-cysteine; hemediol-L-aspartyl ester-L-glutamyl ester; 30 hemediol-L-aspartyl ester-L-glutamyl ester-L-methionine sulfonium; heme-L-cysteine; heme-L-histidine; heparan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-serine; heme P450-bis-L-cysteine-L-lysine; hexakis-L-cysteinyl hexairon hexasulfide; keratan sulfate D-glucuronyl-D-galactosyl-D-galactosyl-D-xylosyl-L-threonine; L

oxoalanine- lactic acid; L phenyllactic acid; l'-(8alpha-FAD)-L-histidine; L-2'.4',5'-topaquinone; L-3',4'-dihydroxyphenylalanine; L-3'.4'.5'-trihydroxyphenylalanine; L-4'-bromophenylalanine; L-6'-bromotryptophan; L-alanine amide; L-alanyl imidazolinone glycine; L-allysine; L-arginine amide; L-asparagine amide; L-aspartic 4-phosphoric anhydride; L-aspartic acid 1-amide; L-beta-methylthioaspartic acid; L-bromohistidine; L-citrulline; L-cysteine amide; L-cysteine glutathione disulfide; L-cysteine methyl disulfide; L-cysteine methyl ester; L-cysteine oxazolecarboxylic acid; L-cysteine oxazolinecarboxylic acid; L-cysteine persulfide; L-cysteine sulfenic acid; L-cysteine sulfenic acid; L-cysteine thiazolecarboxylic acid; L-cysteinyl homocitryl molybdenum-5-heptaiiron-nonasulfide; L-cysteinyl imidazolinone glycine; L-cysteinyl molybdopterin; L-cysteinyl molybdopterin guanine dinucleotide; L-cystine; L-erythro-beta-hydroxyasparagine; L-erythro-beta-hydroxyaspartic acid; L-gamma-carboxyglutamic acid; L-glutamic acid 1-amide; L-glutamic acid 5-methyl ester; L-glutamine amide; L-glutamyl 5-glycerylphosphorylethanolamine; L-histidine amide; L-isoglutamyl-polyglutamic acid; L-isoglutamyl-polyglycine; L-isoleucine amide; L-lanthionine; L-leucine amide; L-lysine amide; L-lysine thiazolecarboxylic acid; L-lysinoalanine; L-methionine amide; L-methionine sulfone; L-phenylalanine thiazolecarboxylic acid; L-phenylalanine amide; L-proline amide; L-selenocysteine; L-selenocysteinyl molybdopterin guanine dinucleotide; L-serine amide; L-serine thiazolecarboxylic acid; L-seryl imidazolinone glycine; L-T-bromophenylalanine; L-T-bromophenylalanine; L-threonine amide; L-thyroxine; L-tryptophan amide; L-tryptophyl quinone; L-tyrosine amide; L-valine amide; meso-lanthionine; N-(L-glutamyl)-L-tyrosine; N-(L-isoaspartyl)-glycine; N-(L-isoaspartyl)-L-cysteine; N,N,N-trimethyl-L-alanine; N,N-dimethyl-L-proline; N2-acetyl-L-lysine; N2-succinyl-L-tryptophan; N4-(ADP-ribosyl)-L-asparagine; N4-glycosyl-L-asparagine; N4-hydroxymethyl-L-asparagine; N4-methyl-L-asparagine; N5-methyl-L-glutamine; N6-1-carboxyethyl-L-lysine; N6-(4-amino hydroxybutyl)-L-lysine; N6-(L-isoglutamyl)-L-lysine; N6-(phospho-5'-adenosine)-L-lysine; N6-(phospho-5'-guanosine)-L-lysine; N6,N6,N6-trimethyl-L-lysine; N6,N6-dimethyl-L-lysine; N6-acetyl-L-lysine; N6-biotinyl-L-lysine; N6-carboxy-L-lysine; N6-formyl-L-lysine; N6-glycyl-L-lysine; N6-lipoyl-L-lysine; N6-methyl-L-lysine; N6-methyl-N6-poly(N-methyl-propylamine)-L-lysine; N6-mureinyl-L-lysine; N6-myristoyl-L-lysine; N6-palmitoyl-L-lysine; N6-pyridoxal phosphate-L-lysine; N6-pyruvic acid 2-iminyl-L-lysine; N6-retinal-L-lysine; N-acetylglycine; N-acetyl-L-glutamine; N-acetyl-L-alanine; N-acetyl-L-aspartic acid; N-

acetyl-L-cysteine; N-acetyl-L-glutamic acid; N-acetyl-L-isoleucine; N-acetyl-L-methionine; N-acetyl-L-proline; N-acetyl-L-serine; N-acetyl-L-threonine; N-acetyl-L-tyrosine; N-acetyl-L-valine; N-alanyl-glycosylphosphatidylinositolethanolamine; N-asparaginyl-glycosylphosphatidylinositolethanolarnine; N-aspartyl-

5 glycosylphosphatidylinositolethanolamine; N-formylglycine; N-formyl-L-methionine; N-glycyl-glycosylphosphatidylinositolethanolamine; N-L-glutamyl-poly-L-glutamic acid; N-methylglycine; N-methyl-L-alanine; N-methyl-L-methionine; N-methyl-L-phenylalanine; N-myristoyl-glycine; N-palmitoyl-L-cysteine; N-pyruvic acid 2-iminyl-L-cysteine; N-pyruvic acid 2-iminyl-L-valine; N-seryl-glycosylphosphatidylinositolethanolamine; N-

10 seryl-glycosyBSPingolipidinositolethanolamine; O-(ADP-ribosyl)-L-serine; O-(phospho-5'-adenosine)-L-threonine; O-(phospho-5'-DNA)-L-serine; O-(phospho-5'-DNA)-L-threonine; O-(phospho-5'rRNA)-L-serine; O-(phosphoribosyl dephospho-coenzyme A)-L-serine; O-(sn-l-glycerophosphoryl)-L-serine; O4'-(8alpha-FAD)-L-tyrosine; O4'-(phospho-5'-adenosine)-L-tyrosine; O4'-(phospho-5'-DNA)-L-tyrosine; O4'-(phospho-5'-RNA)-L-

15 tyrosine; O4'-(phospho-5'-uridine)-L-tyrosine; O4-glycosyl-L-hydroxyproline; O4'-glycosyl-L-tyrosine; O4'-sulfo-L-tyrosine; O5-glycosyl-L-hydroxylysine; O-glycosyl-L-serine; O-glycosyl-L-threonine; omega-N-(ADP-ribosyl)-L-arginine; omega-N-omega-N'-dimethyl-L-arginine; omega-N-methyl-L-arginine; omega-N-omega-N-dimethyl-L-arginine; omega-N-phospho-L-arginine; O'octanoyl-L-serine; O-palmitoyl-L-serine; O-

20 palmitoyl-L-threonine; O-phospho-L-serine; O-phospho-L-threonine; O-phosphopantetheine-L-serine; phyoerythrobilin-bis-L-cysteine; phycourobilin-bis-L-cysteine; pyrroloquinoline quinone; pyruvic acid; S hydroxycinnamyl-L-cysteine; S-(2-aminovinyl) methyl-D-eysteine; S-(2-aminovinyl)-D-cysteine; S-(6-FW-L-cysteine; S-(8alpha-FAD)-L-cysteine; S-(ADP-ribosyl)-L-cysteine; S-(L-isoglutamyl)-L-cysteine; S-

25 12-hydroxyfarnesyl-L-cysteine; S-acetyl-L-cysteine; S-diacylglycerol-L-cysteine; S-diphytanylglycerol diether-L-cysteine; S-farnesyl-L-cysteine; S-geranylgeranyl-L-cysteine; S-glycosyl-L-cysteine; S-glycyl-L-cysteine; S-methyl-L-cysteine; S-nitrosyl-L-cysteine; S-palmitoyl-L-cysteine; S-phospho-L-cysteine; S-phycobiliviolin-L-cysteine; S-phycocyanobilin-L-cysteine; S-phycoerythrobilin-L-cysteine; S-phytochromobilin-L-

30 cysteine; S-selenyl-L-cysteine; S-sulfo-L-cysteine; tetrakis-L-cysteinyl diiron disulfide; tetrakis-L-cysteinyl iron; tetrakis-L-cysteinyl tetrairon tetrasulfide; trans-2,3-cis 4-dihydroxy-L-proline; tris-L-cysteinyl triiron tetrasulfide; tris-L-cysteinyl triiron trisulfide; tris-L-cysteinyl-L-aspartato tetrairon tetrasulfide; tris-L-cysteinyl-L-cysteine persulfido-

bis-L-glutamato-L-histidino tetrairon disulfide trioxide; tris-L-cysteinyl-L-N3'-histidino tetrairon tetrasulfide; tris-L-cysteinyl-L-N1'-histidino tetrairon tetrasulfide; and tris-L-cysteinyl-L-serinyl tetrairon tetrasulfide.

Additional examples of PTMs may be found in web sites such as the Delta Mass 5 database based on Krishna, R. G. and F. Wold (1998). Posttranslational Modifications. Proteins - Analysis and Design. R. H. Angeletti. San Diego, Academic Press. 1: 121-206; Methods in Enzymology, 193, J.A. McClosky (ed) (1990), pages 647-660; Methods in Protein Sequence Analysis edited by Kazutomo Imahori and Fumio Sakiyama, Plenum Press, (1993) "Post-translational modifications of proteins" R.G. Krishna and F. Wold 10 pages 167-172; "GlycoSuiteDB: a new curated relational database of glycoprotein glycan structures and their biological sources" Cooper et al. Nucleic Acids Res. 29; 332-335 (2001) "O-GLYCBASE version 4.0: a revised database of O-glycosylated proteins" Gupta et al. Nucleic Acids Research, 27: 370-372 (1999); and "PhosphoBase, a database of phosphorylation sites: release 2.0.", Kreegipuu et al. Nucleic Acids Res 27(1):237-239 15 (1999) see also, WO 02/21139A2, the disclosure of which is incorporated herein by reference in its entirety.

Tumorigenesis is often accompanied by alterations in the post-translational modifications of proteins. Thus, in another embodiment, the invention provides polypeptides from cancerous cells or tissues that have altered post-translational 20 modifications compared to the post-translational modifications of polypeptides from normal cells or tissues. A number of altered post-translational modifications are known. One common alteration is a change in phosphorylation state, wherein the polypeptide from the cancerous cell or tissue is hyperphosphorylated or hypophosphorylated compared to the polypeptide from a normal tissue, or wherein the polypeptide is phosphorylated on 25 different residues than the polypeptide from a normal cell. Another common alteration is a change in glycosylation state, wherein the polypeptide from the cancerous cell or tissue has more or less glycosylation than the polypeptide from a normal tissue, and/or wherein the polypeptide from the cancerous cell or tissue has a different type of glycosylation than the polypeptide from a noncancerous cell or tissue. Changes in glycosylation may be 30 critical because carbohydrate-protein and carbohydrate-carbohydrate interactions are important in cancer cell progression, dissemination and invasion. See, e.g., Barchi, *Curr. Pharm. Des.* 6: 485-501 (2000), Verma, *Cancer Biochem. Biophys.* 14: 151-162 (1994) and Dennis et al., *Bioessays* 5: 412-421 (1999).

Another post-translational modification that may be altered in cancer cells is prenylation. Prenylation is the covalent attachment of a hydrophobic prenyl group (either farnesyl or geranylgeranyl) to a polypeptide. Prenylation is required for localizing a protein to a cell membrane and is often required for polypeptide function. For instance, 5 the Ras superfamily of GTPase signalling proteins must be prenylated for function in a cell. See, e.g., Prendergast et al., *Semin. Cancer Biol.* 10: 443-452 (2000) and Khwaja et al., *Lancet* 355: 741-744 (2000).

Other post-translation modifications that may be altered in cancer cells include, without limitation, polypeptide methylation, acetylation, arginylation or racemization of 10 amino acid residues. In these cases, the polypeptide from the cancerous cell may exhibit either increased or decreased amounts of the post-translational modification compared to the corresponding polypeptides from noncancerous cells.

Other polypeptide alterations in cancer cells include abnormal polypeptide cleavage of proteins and aberrant protein-protein interactions. Abnormal polypeptide 15 cleavage may be cleavage of a polypeptide in a cancerous cell that does not usually occur in a normal cell, or a lack of cleavage in a cancerous cell, wherein the polypeptide is cleaved in a normal cell. Aberrant protein-protein interactions may be either covalent cross-linking or non-covalent binding between proteins that do not normally bind to each other. Alternatively, in a cancerous cell, a protein may fail to bind to another protein to 20 which it is bound in a noncancerous cell. Alterations in cleavage or in protein-protein interactions may be due to over- or underproduction of a polypeptide in a cancerous cell compared to that in a normal cell, or may be due to alterations in post-translational modifications (see above) of one or more proteins in the cancerous cell. See, e.g., Henschen-Edman, *Ann. N.Y. Acad. Sci.* 936: 580-593 (2001).

25 Alterations in polypeptide post-translational modifications, as well as changes in polypeptide cleavage and protein-protein interactions, may be determined by any method known in the art. For instance, alterations in phosphorylation may be determined by using anti-phosphoserine, anti-phosphothreonine or anti-phosphotyrosine antibodies or by amino acid analysis. Glycosylation alterations may be determined using antibodies specific for 30 different sugar residues, by carbohydrate sequencing, or by alterations in the size of the glycoprotein, which can be determined by, e.g., SDS polyacrylamide gel electrophoresis (PAGE). Other alterations of post-translational modifications, such as prenylation, racemization, methylation, acetylation and arginylation, may be determined by chemical

- analysis, protein sequencing, amino acid analysis, or by using antibodies specific for the particular post-translational modifications. Changes in protein-protein interactions and in polypeptide cleavage may be analyzed by any method known in the art including, without limitation, non-denaturing PAGE (for non-covalent protein-protein interactions), SDS
- 5 PAGE (for covalent protein-protein interactions and protein cleavage), chemical cleavage, protein sequencing or immunoassays.

In another embodiment, the invention provides polypeptides that have been post-translationally modified. In one embodiment, polypeptides may be modified enzymatically or chemically, by addition or removal of a post-translational modification.

10 For example, a polypeptide may be glycosylated or deglycosylated enzymatically. Similarly, polypeptides may be phosphorylated using a purified kinase, such as a MAP kinase (e.g., p38, ERK, or JNK) or a tyrosine kinase (e.g., Src or erbB2). A polypeptide may also be modified through synthetic chemistry. Alternatively, one may isolate the polypeptide of interest from a cell or tissue that expresses the polypeptide with the desired

15 post-translational modification. In another embodiment, a nucleic acid molecule encoding the polypeptide of interest is introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide in the desired fashion. If the polypeptide does not contain a motif for a desired post-translational modification, one may alter the post-translational modification by mutating the nucleic acid sequence of a nucleic

20 acid molecule encoding the polypeptide so that it contains a site for the desired post-translational modification. Amino acid sequences that may be post-translationally modified are known in the art. See, e.g., the programs described above on the website expasy.org of the world wide web. The nucleic acid molecule may also be introduced into a host cell that is capable of post-translationally modifying the encoded polypeptide.

25 Similarly, one may delete sites that are post-translationally modified by either mutating the nucleic acid sequence so that the encoded polypeptide does not contain the post-translational modification motif, or by introducing the native nucleic acid molecule into a host cell that is not capable of post-translationally modifying the encoded polypeptide.

It will be appreciated, as is well known and as noted above, that polypeptides are

30 not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslational events, including natural processing events and events brought about by human manipulation which do not occur naturally. Circular, branched and branched

circular polypeptides may be synthesized by non-translation natural processes and by entirely synthetic methods, as well. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a 5 covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost invariably will be N-formylmethionine.

Useful post-synthetic (and post-translational) modifications include conjugation to 10 detectable labels, such as fluorophores. A wide variety of amine-reactive and thiol-reactive fluorophore derivatives have been synthesized that react under nondenaturing conditions with N-terminal amino groups and epsilon amino groups of lysine residues, on the one hand, and with free thiol groups of cysteine residues, on the other.

Kits are available commercially that permit conjugation of proteins to a variety of 15 amine-reactive or thiol-reactive fluorophores: Molecular Probes, Inc. (Eugene, OR, USA), e.g., offers kits for conjugating proteins to Alexa Fluor 350, Alexa Fluor 430, Fluorescein-EX, Alexa Fluor 488, Oregon Green 488, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 546, Alexa Fluor 568, Alexa Fluor 594, and Texas Red-X.

A wide variety of other amine-reactive and thiol-reactive fluorophores are 20 available commercially (Molecular Probes, Inc., Eugene, OR, USA), including Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, 25 BODIPY 558/568, BODIPY 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA).

The polypeptides of the present invention can also be conjugated to fluorophores, 30 other proteins, and other macromolecules, using bifunctional linking reagents. Common homobifunctional reagents include, e.g., APG, AEDP, BASED, BMB, BMDB, BMH, BMOE, BM[PEO]3, BM[PEO]4, BS3, BSOCOES, DF DN B, DMA, DMP, DMS, DPDPB,

DSG, DSP (Lomant's Reagent), DSS, DST, DTBP, DTME, DTSSP, EGS, HBVS, Sulfo-BSOCOES, Sulfo-DST, Sulfo-EGS (all available from Pierce, Rockford, IL, USA); common heterobifunctional cross-linkers include ABH, AMAS, ANB-NOS, APDP, ASBA, BMPA, BMPH, BMPS, EDC, EMCA, EMCH, EMCS, KMUA, KMUH, GMBS,

- 5 LC-SMCC, LC-SPDP, MBS, M2C2H, MPBH, MSA, NHS-ASA, PDPH, PMPI, SADP, SAED, SAND, SANPAH, SASD, SATP, SBAP, SFAD, SIA, SIAB, SMCC, SMPB, SMPH, SMPT, SPDP, Sulfo-EMCS, Sulfo-GMBS, Sulfo-HSAB, Sulfo-KMUS, Sulfo-LC-SPDP, Sulfo-MBS, Sulfo-NHS-LC-ASA, Sulfo-SADP, Sulfo-SANPAH, Sulfo-SIAB, Sulfo-SMCC, Sulfo-SMPB, Sulfo-LC-SMPT, SVSB, TFCS (all available
- 10 Pierce, Rockford, IL, USA).

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can be conjugated, using such cross-linking reagents, to fluorophores that are not amine- or thiol-reactive. Other labels that usefully can be conjugated to polypeptides of the present invention include radioactive labels,

- 15 echosonographic contrast reagents, and MRI contrast agents.

Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated using cross-linking agents to carrier proteins, such as KLH, bovine thyroglobulin, and even bovine serum albumin (BSA), to increase immunogenicity for raising anti-BSP antibodies.

- 20 Polypeptides of the present invention, including full length polypeptides, fragments and fusion proteins, can also usefully be conjugated to polyethylene glycol (PEG); PEGylation increases the serum half life of proteins administered intravenously for replacement therapy. Delgado *et al.*, *Crit. Rev. Ther. Drug Carrier Syst.* 9(3-4): 249-304 (1992); Scott *et al.*, *Curr. Pharm. Des.* 4(6): 423-38 (1998); De Santis *et al.*, *Curr. Opin. Biotechnol.* 10(4): 324-30 (1999). PEG monomers can be attached to the protein directly or through a linker, with PEGylation using PEG monomers activated with tresyl chloride (2,2,2-trifluoroethanesulphonyl chloride) permitting direct attachment under mild conditions.

- 25 Polypeptides of the present invention are also inclusive of analogs of a polypeptide encoded by a nucleic acid molecule according to the instant invention. In a preferred embodiment, this polypeptide is a BSP. In a more preferred embodiment, this polypeptide is derived from a polypeptide having part or all of the amino acid sequence of SEQ ID NO: 96-232. Also preferred is an analog polypeptide comprising one or more

substitutions of non-natural amino acids or non-native inter-residue bonds compared to the naturally occurring polypeptide. In one embodiment, the analog is structurally similar to a BSP, but one or more peptide linkages is replaced by a linkage selected from the group consisting of --CH₂NH--, --CH₂S--, --CH₂-CH₂--, --CH=CH--(cis and trans), --COCH₂--,

5 --CH(OH)CH₂-- and --CH₂SO--. In another embodiment, the analog comprises substitution of one or more amino acids of a BSP with a D-amino acid of the same type or other non-natural amino acid in order to generate more stable peptides. D-amino acids can readily be incorporated during chemical peptide synthesis: peptides assembled from D-amino acids are more resistant to proteolytic attack; incorporation of D-amino acids can

10 also be used to confer specific three-dimensional conformations on the peptide. Other amino acid analogues commonly added during chemical synthesis include ornithine, norleucine, phosphorylated amino acids (typically phosphoserine, phosphothreonine, phosphotyrosine), L-malonyltyrosine, a non-hydrolyzable analog of phosphotyrosine (*see, e.g.*, Kole *et al.*, *Biochem. Biophys. Res. Com.* 209: 817-821 (1995)), and various

15 halogenated phenylalanine derivatives.

Non-natural amino acids can be incorporated during solid phase chemical synthesis or by recombinant techniques, although the former is typically more common. Solid phase chemical synthesis of peptides is well established in the art. Procedures are described, *inter alia*, in Chan *et al.* (eds.), Fmoc Solid Phase Peptide Synthesis: A Practical Approach (Practical Approach Series), Oxford Univ. Press (March 2000); Jones, Amino Acid and Peptide Synthesis (Oxford Chemistry Primers, No 7), Oxford Univ. Press (1992); and Bodanszky, Principles of Peptide Synthesis (Springer Laboratory), Springer Verlag (1993).

Amino acid analogues having detectable labels are also usefully incorporated

25 during synthesis to provide derivatives and analogs. Biotin, for example can be added using biotinoyl-(9-fluorenylmethoxycarbonyl)-L-lysine (FMOC biocytin) (Molecular Probes, Eugene, OR, USA). Biotin can also be added enzymatically by incorporation into a fusion protein of an *E. coli* BirA substrate peptide. The FMOC and tBOC derivatives of dabcyl-L-lysine (Molecular Probes, Inc., Eugene, OR, USA) can be used to incorporate

30 the dabcyl chromophore at selected sites in the peptide sequence during synthesis. The aminonaphthalene derivative EDANS, the most common fluorophore for pairing with the dabcyl quencher in fluorescence resonance energy transfer (FRET) systems, can be

introduced during automated synthesis of peptides by using EDANS-FMOC-L-glutamic acid or the corresponding *t*BOC derivative (both from Molecular Probes, Inc., Eugene, OR, USA). Tetramethylrhodamine fluorophores can be incorporated during automated FMOC synthesis of peptides using (FMOC)-TMR-L-lysine (Molecular Probes, Inc. 5 Eugene, OR, USA).

Other useful amino acid analogues that can be incorporated during chemical synthesis include aspartic acid, glutamic acid, lysine, and tyrosine analogues having allyl side-chain protection (Applied Biosystems, Inc., Foster City, CA, USA); the allyl side chain permits synthesis of cyclic, branched-chain, sulfonated, glycosylated, and 10 phosphorylated peptides.

A large number of other FMOC-protected non-natural amino acid analogues capable of incorporation during chemical synthesis are available commercially, including, e.g., Fmoc-2-aminobicyclo[2.2.1]heptane-2-carboxylic acid, Fmoc-3-endo-aminobicyclo[2.2.1]heptane-2-endo-carboxylic acid, Fmoc-3-exo-aminobicyclo[2.2.1]heptane-2-exo-carboxylic acid, Fmoc-3-endo-amino-bicyclo[2.2.1]hept-5-ene-2-endo-carboxylic acid, Fmoc-3-exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid, Fmoc-cis-2-amino-1-cyclohexanecarboxylic acid, Fmoc-trans-2-amino-1-cyclohexanecarboxylic acid, Fmoc-1-amino-1-cyclopentanecarboxylic acid, Fmoc-cis-2-amino-1-cyclopentanecarboxylic acid, Fmoc-1-amino-1-cyclopropanecarboxylic acid, Fmoc-D-2-amino-4-(ethylthio)butyric acid, Fmoc-L-2-amino-4-(ethylthio)butyric acid, Fmoc-L-buthionine, Fmoc-S-methyl-L-Cysteine, Fmoc-2-aminobenzoic acid (anthranillic acid), Fmoc-3-aminobenzoic acid, Fmoc-4-aminobenzoic acid, Fmoc-2-aminobenzophenone-2'-carboxylic acid, Fmoc-N-(4-aminobenzoyl)- β -alanine, Fmoc-2-amino-4,5-dimethoxybenzoic acid, Fmoc-4-aminohippuric acid, Fmoc-2-amino-3-hydroxybenzoic acid, Fmoc-2-amino-5-hydroxybenzoic acid, Fmoc-3-amino-4-hydroxybenzoic acid, Fmoc-4-amino-3-hydroxybenzoic acid, Fmoc-4-amino-2-hydroxybenzoic acid, Fmoc-5-amino-2-hydroxybenzoic acid, Fmoc-2-amino-3-methoxybenzoic acid, Fmoc-4-amino-3-methoxybenzoic acid, Fmoc-2-amino-3-methylbenzoic acid, Fmoc-2-amino-5-methylbenzoic acid, Fmoc-2-amino-6-methylbenzoic acid, Fmoc-3-amino-2-methylbenzoic acid, Fmoc-3-amino-4-methylbenzoic acid, Fmoc-4-amino-3-methylbenzoic acid, Fmoc-3-amino-2-naphtoic acid, Fmoc-D,L-3-amino-3-phenylpropionic acid, Fmoc-L-Methyldopa, Fmoc-2-amino-4,6-dimethyl-3-

pyridinecarboxylic acid, Fmoc-D,L-amino-2-thiophenacetic acid, Fmoc-4-(carboxymethyl)piperazine, Fmoc-4-carboxypiperazine, Fmoc-4-(carboxymethyl)homopiperazine, Fmoc-4-phenyl-4-piperidinecarboxylic acid, Fmoc-L-1,2,3,4-tetrahydronorharman-3-carboxylic acid, Fmoc-L-thiazolidine-4-carboxylic acid, all 5 available from The Peptide Laboratory (Richmond, CA, USA).

Non-natural residues can also be added biosynthetically by engineering a suppressor tRNA, typically one that recognizes the UAG stop codon, by chemical aminoacylation with the desired unnatural amino acid. Conventional site-directed mutagenesis is used to introduce the chosen stop codon UAG at the site of interest in the 10 protein gene. When the acylated suppressor tRNA and the mutant gene are combined in an *in vitro* transcription/translation system, the unnatural amino acid is incorporated in response to the UAG codon to give a protein containing that amino acid at the specified position. Liu *et al.*, *Proc. Natl Acad. Sci. USA* 96(9): 4780-5 (1999); Wang *et al.*, *Science* 292(5516): 498-500 (2001).

15 *Fusion Proteins*

Another aspect of the present invention relates to the fusion of a polypeptide of the present invention to heterologous polypeptides. In a preferred embodiment, the polypeptide of the present invention is a BSP. In a more preferred embodiment, the polypeptide of the present invention that is fused to a heterologous polypeptide which 20 comprises part or all of the amino acid sequence of SEQ ID NO: 96-232, or is a mutein, homologous polypeptide, analog or derivative thereof. In an even more preferred embodiment, the fusion protein is encoded by a nucleic acid molecule comprising all or part of the nucleic acid sequence of SEQ ID NO: 1-95, or comprises all or part of a nucleic acid sequence that selectively hybridizes or is homologous to a nucleic acid molecule 25 comprising a nucleic acid sequence of SEQ ID NO: 1-95.

The fusion proteins of the present invention will include at least one fragment of a polypeptide of the present invention, which fragment is at least 6, typically at least 8, often at least 15, and usefully at least 16, 17, 18, 19, or 20 amino acids long. The fragment of the polypeptide of the present to be included in the fusion can usefully be at least 25 30 amino acids long, at least 50 amino acids long, and can be at least 75, 100, or even 150 amino acids long. Fusions that include the entirety of a polypeptide of the present invention have particular utility.

The heterologous polypeptide included within the fusion protein of the present invention is at least 6 amino acids in length, often at least 8 amino acids in length, and preferably at least 15, 20, or 25 amino acids in length. Fusions that include larger polypeptides, such as the IgG Fc region, and even entire proteins (such as GFP 5 chromophore-containing proteins) are particularly useful.

As described above in the description of vectors and expression vectors of the present invention, which discussion is incorporated here by reference in its entirety, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those designed to facilitate purification and/or visualization of 10 recombinantly-expressed proteins. *See, e.g., Ausubel, Chapter 16, (1992), supra.* Although purification tags can also be incorporated into fusions that are chemically synthesized, chemical synthesis typically provides sufficient purity that further purification by HPLC suffices; however, visualization tags as above described retain their utility even when the protein is produced by chemical synthesis, and when so included 15 render the fusion proteins of the present invention useful as directly detectable markers of the presence of a polypeptide of the invention.

As also discussed above, heterologous polypeptides to be included in the fusion proteins of the present invention can usefully include those that facilitate secretion of recombinantly expressed proteins into the periplasmic space or extracellular milieu for 20 prokaryotic hosts or into the culture medium for eukaryotic cells through incorporation of secretion signals and/or leader sequences. For example, a His⁶ tagged protein can be purified on a Ni affinity column and a GST fusion protein can be purified on a glutathione affinity column. Similarly, a fusion protein comprising the Fc domain of IgG can be purified on a Protein A or Protein G column and a fusion protein comprising an epitope 25 tag such as myc can be purified using an immunoaffinity column containing an anti-c-myc antibody. It is preferable that the epitope tag be separated from the protein encoded by the essential gene by an enzymatic cleavage site that can be cleaved after purification. See also the discussion of nucleic acid molecules encoding fusion proteins that may be expressed on the surface of a cell.

30 Other useful fusion proteins of the present invention include those that permit use of the polypeptide of the present invention as bait in a yeast two-hybrid system. *See Bartel et al. (eds.), The Yeast Two-Hybrid System, Oxford University Press (1997); Zhu et al., Yeast Hybrid Technologies, Eaton Publishing (2000); Fields et al., Trends Genet.*

10(8): 286-92 (1994); Mendelsohn *et al.*, *Curr. Opin. Biotechnol.* 5(5): 482-6 (1994); Luban *et al.*, *Curr. Opin. Biotechnol.* 6(1): 59-64 (1995); Allen *et al.*, *Trends Biochem. Sci.* 20(12): 511-6 (1995); Drees, *Curr. Opin. Chem. Biol.* 3(1): 64-70 (1999); Topcu *et al.*, *Pharm. Res.* 17(9): 1049-55 (2000); Fashena *et al.*, *Gene* 250(1-2): 1-14 (2000); Colas 5 *et al.*, *Nature* 380, 548-550 (1996); Norman, T. *et al.*, *Science* 285, 591-595 (1999); Fabbrizio *et al.*, *Oncogene* 18, 4357-4363 (1999); Xu *et al.*, *Proc Natl Acad Sci U S A.* 94, 12473-12478 (1997); Yang, *et al.*, *Nuc. Acids Res.* 23, 1152-1156 (1995); Kolonin *et al.*, *Proc Natl Acad Sci U S A* 95, 14266-14271 (1998); Cohen *et al.*, *Proc Natl Acad Sci U S A* 95, 14272-14277 (1998); Uetz, *et al.* *Nature* 403, 623-627(2000); Ito, *et al.*, *Proc Natl Acad Sci U S A* 98, 4569-4574 (2001). Typically, such fusion is to either *E. coli* LexA or yeast GAL4 DNA binding domains. Related bait plasmids are available that express the bait fused to a nuclear localization signal.

Other useful fusion proteins include those that permit display of the encoded polypeptide on the surface of a phage or cell, fusions to intrinsically fluorescent proteins, 15 such as green fluorescent protein (GFP), and fusions to the IgG Fc region, as described above.

The polypeptides of the present invention can also usefully be fused to protein toxins, such as *Pseudomonas* exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin, in order to effect ablation of cells that bind or take up the proteins of 20 the present invention.

Fusion partners include, *inter alia*, *myc*, hemagglutinin (HA), GST, immunoglobulins, β -galactosidase, biotin trpE, protein A, β -lactamase, α -amylase, maltose binding protein, alcohol dehydrogenase, polyhistidine (for example, six histidine at the amino and/or carboxyl terminus of the polypeptide), lacZ, green fluorescent protein 25 (GFP), yeast α mating factor, GAL4 transcription activation or DNA binding domain, luciferase, and serum proteins such as ovalbumin, albumin and the constant domain of IgG. See, e.g., Ausubel (1992), *supra* and Ausubel (1999), *supra*. Fusion proteins may also contain sites for specific enzymatic cleavage, such as a site that is recognized by enzymes such as Factor XIII, trypsin, pepsin, or any other enzyme known in the art. 30 Fusion proteins will typically be made by either recombinant nucleic acid methods, as described above, chemically synthesized using techniques well known in the art (e.g., a Merrifield synthesis), or produced by chemical cross-linking.

Another advantage of fusion proteins is that the epitope tag can be used to bind the fusion protein to a plate or column through an affinity linkage for screening binding proteins or other molecules that bind to the BSP.

- As further described below, the polypeptides of the present invention can readily
- 5 be used as specific immunogens to raise antibodies that specifically recognize polypeptides of the present invention including BSPs and their allelic variants and homologues. The antibodies, in turn, can be used, *inter alia*, specifically to assay for the polypeptides of the present invention, particularly BSPs, *e.g.* by ELISA for detection of protein fluid samples, such as serum, by immunohistochemistry or laser scanning
- 10 cytometry, for detection of protein in tissue samples, or by flow cytometry, for detection of intracellular protein in cell suspensions, for specific antibody-mediated isolation and/or purification of BSPs, as for example by immunoprecipitation, and for use as specific agonists or antagonists of BSPs.

- One may determine whether polypeptides of the present invention including BSPs,
- 15 muteins, homologous proteins or allelic variants or fusion proteins of the present invention are functional by methods known in the art. For instance, residues that are tolerant of change while retaining function can be identified by altering the polypeptide at known residues using methods known in the art, such as alanine scanning mutagenesis, Cunningham *et al.*, *Science* 244(4908): 1081-5 (1989); transposon linker scanning
- 20 mutagenesis, Chen *et al.*, *Gene* 263(1-2): 39-48 (2001); combinations of homolog- and alanine-scanning mutagenesis, Jin *et al.*, *J. Mol. Biol.* 226(3): 851-65 (1992); and combinatorial alanine scanning, Weiss *et al.*, *Proc. Natl. Acad. Sci USA* 97(16): 8950-4 (2000), followed by functional assay. Transposon linker scanning kits are available commercially (New England Biolabs, Beverly, MA, USA, catalog. no. E7-102S;
- 25 EZ::TNTTM In-Frame Linker Insertion Kit, catalogue no. EZI04KN, (Epicentre Technologies Corporation, Madison, WI, USA).

- Purification of the polypeptides or fusion proteins of the present invention is well known and within the skill of one having ordinary skill in the art. *See, e.g.*, Scopes, Protein Purification, 2d ed. (1987). Purification of recombinantly expressed polypeptides
- 30 is described above. Purification of chemically-synthesized peptides can readily be effected, *e.g.*, by HPLC.

Accordingly, it is an aspect of the present invention to provide the isolated polypeptides or fusion proteins of the present invention in pure or substantially pure form

in the presence or absence of a stabilizing agent. Stabilizing agents include both proteinaceous and non-proteinaceous material and are well known in the art. Stabilizing agents, such as albumin and polyethylene glycol (PEG) are known and are commercially available.

5 Although high levels of purity are preferred when the isolated polypeptide or fusion protein of the present invention are used as therapeutic agents, such as in vaccines and replacement therapy, the isolated polypeptides of the present invention are also useful at lower purity. For example, partially purified polypeptides of the present invention can be used as immunogens to raise antibodies in laboratory animals.

10 In a preferred embodiment, the purified and substantially purified polypeptides of the present invention are in compositions that lack detectable ampholytes, acrylamide monomers, bis-acrylamide monomers, and polyacrylamide.

15 The polypeptides or fusion proteins of the present invention can usefully be attached to a substrate. The substrate can be porous or solid, planar or non-planar; the bond can be covalent or noncovalent. For example, the peptides of the invention may be stabilized by covalent linkage to albumin. See, U.S. Patent No. 5,876,969, the contents of which are hereby incorporated in its entirety.

20 The polypeptides or fusion proteins of the present invention can also be usefully bound to a porous substrate, commonly a membrane, typically comprising nitrocellulose, polyvinylidene fluoride (PVDF), or cationically derivatized, hydrophilic PVDF; so bound, the polypeptides or fusion proteins of the present invention can be used to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized polypeptide or fusion protein of the present invention.

25 As another example, the polypeptides or fusion proteins of the present invention can usefully be bound to a substantially nonporous substrate, such as plastic, to detect and quantify antibodies, *e.g.* in serum, that bind specifically to the immobilized protein of the present invention. Such plastics include polymethylacrylic, polyethylene, polypropylene, polyacrylate, polymethylmethacrylate, polyvinylchloride, polytetrafluoroethylene, polystyrene, polycarbonate, polyacetal, polysulfone, celluloseacetate, cellulosenitrate, 30 nitrocellulose, or mixtures thereof; when the assay is performed in a standard microtiter dish, the plastic is typically polystyrene.

The polypeptides and fusion proteins of the present invention can also be attached to a substrate suitable for use as a surface enhanced laser desorption ionization source; so

attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biologic interaction there between. The polypeptides or fusion proteins of the present invention can also be attached to a substrate suitable for use in surface plasmon resonance detection; so attached, the polypeptide or fusion protein of the present invention is useful for binding and then detecting secondary proteins that bind with sufficient affinity or avidity to the surface-bound polypeptide or fusion protein to indicate biological interaction there between.

Alternative Transcripts

10 In another aspect, the present invention provides splice variants of genes and proteins encoded thereby. The identification of a novel splice variant which encodes an amino acid sequence with a novel region can be targeted for the generation of reagents for use in detection and/or treatment of cancer. The novel amino acid sequence may lead to a unique protein structure, protein subcellular localization, biochemical processing or 15 function of the splice variant. This information can be used to directly or indirectly facilitate the generation of additional or novel therapeutics or diagnostics. The nucleotide sequence in this novel splice variant can be used as a nucleic acid probe for the diagnosis and/or treatment of cancer.

20 Specifically, the newly identified sequences may enable the production of new antibodies or compounds directed against the novel region for use as a therapeutic or diagnostic. Alternatively, the newly identified sequences may alter the biochemical or biological properties of the encoded protein in such a way as to enable the generation of improved or different therapeutics targeting this protein.

Antibodies

25 In another aspect, the invention provides antibodies, including fragments and derivatives thereof, that bind specifically to polypeptides encoded by the nucleic acid molecules of the invention. In a preferred embodiment, the antibodies are specific for a polypeptide that is a BSP, or a fragment, mutein, derivative, analog or fusion protein thereof. In a more preferred embodiment, the antibodies are specific for a polypeptide that 30 comprises SEQ ID NO: 96-232, or a fragment, mutein, derivative, analog or fusion protein thereof.

The antibodies of the present invention can be specific for linear epitopes, discontinuous epitopes, or conformational epitopes of such proteins or protein fragments, either as present on the protein in its native conformation or, in some cases, as present on the proteins as denatured, as, e.g., by solubilization in SDS. New epitopes may also be
5 due to a difference in post translational modifications (PTMs) in disease versus normal tissue. For example, a particular site on a BSP may be glycosylated in cancerous cells, but not glycosylated in normal cells or vice versa. In addition, alternative splice forms of a
10 BSP may be indicative of cancer. Differential degradation of the C or N-terminus of a BSP may also be a marker or target for anticancer therapy. For example, a BSP may be
15 N-terminal degraded in cancer cells exposing new epitopes to antibodies which may selectively bind for diagnostic or therapeutic uses.

As is well known in the art, the degree to which an antibody can discriminate as among molecular species in a mixture will depend, in part, upon the conformational relatedness of the species in the mixture; typically, the antibodies of the present invention
15 will discriminate over adventitious binding to non-BSP polypeptides by at least two-fold, more typically by at least 5-fold, typically by more than 10-fold, 25-fold, 50-fold, 75-fold, and often by more than 100-fold, and on occasion by more than 500-fold or 1000-fold.
When used to detect the proteins or protein fragments of the present invention, the antibody of the present invention is sufficiently specific when it can be used to determine
20 the presence of the polypeptide of the present invention in samples derived from human breast.

Typically, the affinity or avidity of an antibody (or antibody multimer, as in the case of an IgM pentamer) of the present invention for a protein or protein fragment of the present invention will be at least about 1×10^{-6} molar (M), typically at least about 5×10^{-7}
25 M, 1×10^{-7} M, with affinities and avidities of at least 1×10^{-8} M, 5×10^{-9} M, 1×10^{-10} M and up to 1×10^{-13} M proving especially useful.

The antibodies of the present invention can be naturally occurring forms, such as IgG, IgM, IgD, IgE, IgY, and IgA, from any avian, reptilian, or mammalian species.

Human antibodies can, but will infrequently, be drawn directly from human donors
30 or human cells. In such case, antibodies to the polypeptides of the present invention will typically have resulted from fortuitous immunization, such as autoimmune immunization, with the polypeptide of the present invention. Such antibodies will typically, but will not

invariably, be polyclonal. In addition, individual polyclonal antibodies may be isolated and cloned to generate monoclonals.

Human antibodies are more frequently obtained using transgenic animals that express human immunoglobulin genes, which transgenic animals can be affirmatively 5 immunized with the protein immunogen of the present invention. Human Ig-transgenic mice capable of producing human antibodies and methods of producing human antibodies therefrom upon specific immunization are described, *inter alia*, in U.S. Patent Nos. 6,162,963; 6,150,584; 6,114,598; 6,075,181; 5,939,598; 5,877,397; 5,874,299; 5,814,318; 5,789,650; 5,770,429; 5,661,016; 5,633,425; 5,625,126; 5,569,825; 5,545,807; 5,545,806, 10 and 5,591,669, the disclosures of which are incorporated herein by reference in their entireties. Such antibodies are typically monoclonal, and are typically produced using techniques developed for production of murine antibodies.

Human antibodies are particularly useful, and often preferred, when the antibodies of the present invention are to be administered to human beings as *in vivo* diagnostic or 15 therapeutic agents, since recipient immune response to the administered antibody will often be substantially less than that occasioned by administration of an antibody derived from another species, such as mouse.

IgG, IgM, IgD, IgE, IgY, and IgA antibodies of the present invention are also usefully obtained from other species, including mammals such as rodents (typically 20 mouse, but also rat, guinea pig, and hamster), lagomorphs (typically rabbits), and also larger mammals, such as sheep, goats, cows, and horses; or egg laying birds or reptiles such as chickens or alligators. In such cases, as with the transgenic human-antibody-producing non-human mammals, fortuitous immunization is not required, and the non-human mammal is typically affirmatively immunized, according to standard immunization 25 protocols, with the polypeptide of the present invention. One form of avian antibodies may be generated using techniques described in WO 00/29444, published 25 May 2000, which is herein incorporated by reference in its entirety.

As discussed above, virtually all fragments of 8 or more contiguous amino acids of a polypeptide of the present invention can be used effectively as immunogens when 30 conjugated to a carrier, typically a protein such as bovine thyroglobulin, keyhole limpet hemocyanin, or bovine serum albumin, conveniently using a bifunctional linker such as those described elsewhere above, which discussion is incorporated by reference here.

Immunogenicity can also be conferred by fusion of the polypeptide of the present invention to other moieties. For example, polypeptides of the present invention can be produced by solid phase synthesis on a branched polylysine core matrix; these multiple antigenic peptides (MAPs) provide high purity, increased avidity, accurate chemical definition and improved safety in vaccine development. Tam *et al.*, *Proc. Natl. Acad. Sci. USA* 85: 5409-5413 (1988); Posnett *et al.*, *J. Biol. Chem.* 263: 1719-1725 (1988).

Protocols for immunizing non-human mammals or avian species are well-established in the art. See Harlow *et al.* (eds.), Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory (1998); Coligan *et al.* (eds.), Current Protocols in Immunology, John Wiley & Sons, Inc. (2001); Zola, Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives (Basics: From Background to Bench), Springer Verlag (2000); Gross M, Speck *J.Dtsch. Tierarztl. Wochenschr.* 103: 417-422 (1996). Immunization protocols often include multiple immunizations, either with or without adjuvants such as Freund's complete adjuvant and Freund's incomplete adjuvant, and may include naked DNA immunization. Moss, *Semin. Immunol.* 2: 317-327 (1990).

Antibodies from non-human mammals and avian species can be polyclonal or monoclonal, with polyclonal antibodies having certain advantages in immunohistochemical detection of the polypeptides of the present invention and monoclonal antibodies having advantages in identifying and distinguishing particular epitopes of the polypeptides of the present invention. Antibodies from avian species may have particular advantage in detection of the polypeptides of the present invention, in human serum or tissues. Vikinge *et al.*, *Biosens. Bioelectron.* 13: 1257-1262 (1998). Following immunization, the antibodies of the present invention can be obtained using any art-accepted technique. Such techniques are well known in the art and are described in detail in references such as Coligan, *supra*; Zola, *supra*; Howard *et al.* (eds.), Basic Methods in Antibody Production and Characterization, CRC Press (2000); Harlow, *supra*; Davis (ed.), Monoclonal Antibody Protocols, Vol. 45, Humana Press (1995); Delves (ed.), Antibody Production: Essential Techniques, John Wiley & Son Ltd (1997); and Kenney, *Antibody Solution: An Antibody Methods Manual*, Chapman & Hall (1997).

Briefly, such techniques include, *inter alia*, production of monoclonal antibodies by hybridomas and expression of antibodies or fragments or derivatives thereof from host cells engineered to express immunoglobulin genes or fragments thereof. These two

methods of production are not mutually exclusive: genes encoding antibodies specific for the polypeptides of the present invention can be cloned from hybridomas and thereafter expressed in other host cells. Nor need the two necessarily be performed together: e.g., genes encoding antibodies specific for the polypeptides of the present invention can be 5 cloned directly from B cells known to be specific for the desired protein, as further described in U.S. Patent No. 5,627,052, the disclosure of which is incorporated herein by reference in its entirety, or from antibody-displaying phage.

Recombinant expression in host cells is particularly useful when fragments or derivatives of the antibodies of the present invention are desired.

10 Host cells for recombinant antibody production of whole antibodies, antibody fragments, or antibody derivatives can be prokaryotic or eukaryotic.

Prokaryotic hosts are particularly useful for producing phage displayed antibodies of the present invention.

15 The technology of phage-displayed antibodies, in which antibody variable region fragments are fused, for example, to the gene III protein (pIII) or gene VIII protein (pVIII) for display on the surface of filamentous phage, such as M13, is by now well-established. See, e.g., Sidhu, *Curr. Opin. Biotechnol.* 11(6): 610-6 (2000); Griffiths *et al.*, *Curr. Opin. Biotechnol.* 9(1): 102-8 (1998); Hoogenboom *et al.*, *Immunotechnology*, 4(1): 1-20 (1998); Rader *et al.*, *Current Opinion in Biotechnology* 8: 503-508 (1997); Aujame *et al.*, *Human Antibodies* 8: 155-168 (1997); Hoogenboom, *Trends in Biotechnol.* 15: 62-70 (1997); de Kruif *et al.*, 17: 453-455 (1996); Barbas *et al.*, *Trends in Biotechnol.* 14: 230-234 (1996); Winter *et al.*, *Ann. Rev. Immunol.* 433-455 (1994). Techniques and protocols required to generate, propagate, screen (pan), and use the antibody fragments from such libraries have 20 recently been compiled. See, e.g., Barbas (2001), *supra*; Kay, *supra*; and Abelson, *supra*.

25 Typically, phage-displayed antibody fragments are scFv fragments or Fab fragments; when desired, full length antibodies can be produced by cloning the variable regions from the displaying phage into a complete antibody and expressing the full length antibody in a further prokaryotic or a eukaryotic host cell. Eukaryotic cells are also useful for expression of the antibodies, antibody fragments, and antibody derivatives of the 30 present invention. For example, antibody fragments of the present invention can be produced in *Pichia pastoris* and in *Saccharomyces cerevisiae*. See, e.g., Takahashi *et al.*, *Biosci. Biotechnol. Biochem.* 64(10): 2138-44 (2000); Freyre *et al.*, *J. Biotechnol.* 76(2-3): 1 57-63 (2000); Fischer *et al.*, *Biotechnol. Appl. Biochem.* 30 (Pt 2): 117-20

(1999); Pennell *et al.*, *Res. Immunol.* 149(6): 599-603 (1998); Eldin *et al.*, *J. Immunol. Methods.* 201(1): 67-75 (1997);, Frenken *et al.*, *Res. Immunol.* 149(6): 589-99 (1998); and Shusta *et al.*, *Nature Biotechnol.* 16(8): 773-7 (1998).

Antibodies, including antibody fragments and derivatives, of the present invention
5 can also be produced in insect cells. See, e.g., Li *et al.*, *Protein Expr. Purif.* 21(1): 121-8
(2001); Ailor *et al.*, *Biotechnol. Bioeng.* 58(2-3): 196-203 (1998); Hsu *et al.*, *Biotechnol.*
Prog. 13(1): 96-104 (1997); Edelman *et al.*, *Immunology* 91(1): 13-9 (1997); and Nesbit *et*
al., *J. Immunol. Methods* 151(1-2): 201-8 (1992).

Antibodies and fragments and derivatives thereof of the present invention can also
10 be produced in plant cells, particularly maize or tobacco, Giddings *et al.*, *Nature*
Biotechnol. 18(11): 1151-5 (2000); Gavilondo *et al.*, *Biotechniques* 29(1): 128-38 (2000);
Fischer *et al.*, *J. Biol. Regul. Homeost. Agents* 14(2): 83-92 (2000); Fischer *et al.*,
Biotechnol. Appl. Biochem. 30 (Pt 2): 113-6 (1999); Fischer *et al.*, *Biol. Chem.* 380(7-8):
825-39 (1999); Russell, *Curr. Top. Microbiol. Immunol.* 240: 119-38 (1999); and Ma *et*
15 *al.*, *Plant Physiol.* 109(2): 341-6 (1995).

Antibodies, including antibody fragments and derivatives, of the present invention
can also be produced in transgenic, non-human, mammalian milk. See, e.g. Pollock *et al.*,
J. Immunol Methods. 231: 147-57 (1999); Young *et al.*, *Res. Immunol.* 149: 609-10
(1998); and Limonta *et al.*, *Immunotechnology* 1: 107-13 (1995).

Mammalian cells useful for recombinant expression of antibodies, antibody
20 fragments, and antibody derivatives of the present invention include CHO cells, COS
cells, 293 cells, and myeloma cells. Verma *et al.*, *J. Immunol. Methods* 216(1-2):165-81
(1998) review and compare bacterial, yeast, insect and mammalian expression systems for
expression of antibodies. Antibodies of the present invention can also be prepared by cell
25 free translation, as further described in Merk *et al.*, *J. Biochem.* (Tokyo) 125(2): 328-33
(1999) and Ryabova *et al.*, *Nature Biotechnol.* 15(1): 79-84 (1997), and in the milk of
transgenic animals, as further described in Pollock *et al.*, *J. Immunol. Methods* 231(1-2):
147-57 (1999).

The invention further provides antibody fragments that bind specifically to one or
30 more of the polypeptides of the present invention or to one or more of the polypeptides
encoded by the isolated nucleic acid molecules of the present invention, or the binding of
which can be competitively inhibited by one or more of the polypeptides of the present
invention or one or more of the polypeptides encoded by the isolated nucleic acid

molecules of the present invention. Among such useful fragments are Fab, Fab', Fv, F(ab)'₂, and single-chain Fv (scFv) fragments. Other useful fragments are described in Hudson, *Curr. Opin. Biotechnol.* 9(4): 395-402 (1998).

The present invention also relates to antibody derivatives that bind specifically to
5 one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention.

10 Among such useful derivatives are chimeric, primatized, and humanized antibodies; such derivatives are less immunogenic in human beings, and thus are more suitable for *in vivo* administration, than are unmodified antibodies from non-human mammalian species. Another useful method is PEGylation to increase the serum half life of the antibodies.

15 Chimeric antibodies typically include heavy and/or light chain variable regions (including both CDR and framework residues) of immunoglobulins of one species, typically mouse, fused to constant regions of another species, typically human. See, e.g., Morrison *et al.*, *Proc. Natl. Acad. Sci USA* 81(21): 6851-5 (1984); Sharon *et al.*, *Nature* 309(5966): 364-7 (1984); Takeda *et al.*, *Nature* 314(6010): 452-4 (1985); and U.S. Patent

20 No. 5,807,715 the disclosure of which is incorporated herein by reference in its entirety. Primatized and humanized antibodies typically include heavy and/or light chain CDRs from a murine antibody grafted into a non-human primate or human antibody V region framework, usually further comprising a human constant region, Riechmann *et al.*, *Nature* 332(6162): 323-7 (1988); Co *et al.*, *Nature* 351(6326): 501-2 (1991); and U.S. Patent Nos. 25 6,054,297; 5,821,337; 5,770,196; 5,766,886; 5,821,123; 5,869,619; 6,180,377; 6,013,256; 5,693,761; and 6,180,370, the disclosures of which are incorporated herein by reference in their entireties. Other useful antibody derivatives of the invention include heteromeric antibody complexes and antibody fusions, such as diabodies (bispecific antibodies), single-chain diabodies, and intrabodies.

30 It is contemplated that the nucleic acids encoding the antibodies of the present invention can be operably joined to other nucleic acids forming a recombinant vector for cloning or for expression of the antibodies of the invention. Accordingly, the present invention includes any recombinant vector containing the coding sequences, or part

thereof, whether for eukaryotic transduction, transfection or gene therapy. Such vectors may be prepared using conventional molecular biology techniques, known to those with skill in the art, and would comprise DNA encoding sequences for the immunoglobulin V-regions including framework and CDRs or parts thereof, and a suitable promoter either
5 with or without a signal sequence for intracellular transport. Such vectors may be transduced or transfected into eukaryotic cells or used for gene therapy (Marasco et al., *Proc. Natl. Acad. Sci. (USA)* 90: 7889-7893 (1993); Duan et al., *Proc. Natl. Acad. Sci. (USA)* 91: 5075-5079 (1994), by conventional techniques, known to those with skill in the art.

10 The antibodies of the present invention, including fragments and derivatives thereof, can usefully be labeled. It is, therefore, another aspect of the present invention to provide labeled antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited
15 by one or more of the polypeptides of the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention. The choice of label depends, in part, upon the desired use.

For example, when the antibodies of the present invention are used for immunohistochemical staining of tissue samples, the label can usefully be an enzyme that
20 catalyzes production and local deposition of a detectable product. Enzymes typically conjugated to antibodies to permit their immunohistochemical visualization are well known, and include alkaline phosphatase, β-galactosidase, glucose oxidase, horseradish peroxidase (HRP), and urease. Typical substrates for production and deposition of visually detectable products include o-nitrophenyl-beta-D-galactopyranoside (ONPG);
25 o-phenylenediamine dihydrochloride (OPD); p-nitrophenyl phosphate (PNPP); p-nitrophenyl-beta-D-galactopyranoside (PNPG); 3',3'-diaminobenzidine (DAB); 3-amino-9-ethylcarbazole (AEC); 4-chloro-1-naphthol (CN); 5-bromo-4-chloro-3-indolyl-phosphate (BCIP); ABTS®; BluoGal; iodonitrotetrazolium (INT); nitroblue tetrazolium chloride (NBT); phenazine methosulfate (PMS);
30 phenolphthalein monophosphate (PMP); tetramethyl benzidine (TMB); tetranitroblue tetrazolium (TNBT); X-Gal; X-Gluc; and X-Glucoside.

Other substrates can be used to produce products for local deposition that are luminescent. For example, in the presence of hydrogen peroxide (H₂O₂), horseradish

peroxidase (HRP) can catalyze the oxidation of cyclic diacylhydrazides, such as luminol. Immediately following the oxidation, the luminol is in an excited state (intermediate reaction product), which decays to the ground state by emitting light. Strong enhancement of the light emission is produced by enhancers, such as phenolic compounds. Advantages 5 include high sensitivity, high resolution, and rapid detection without radioactivity and requiring only small amounts of antibody. See, e.g., Thorpe *et al.*, *Methods Enzymol.* 133: 331-53 (1986); Kricka *et al.*, *J. Immunoassay* 17(1): 67-83 (1996); and Lundqvist *et al.*, *J. Biolumin. Chemilumin.* 10(6): 353-9 (1995). Kits for such enhanced chemiluminescent detection (ECL) are available commercially. The antibodies can also be labeled using 10 colloidal gold.

As another example, when the antibodies of the present invention are used, e.g., for flow cytometric detection, for scanning laser cytometric detection, or for fluorescent immunoassay, they can usefully be labeled with fluorophores. There are a wide variety of fluorophore labels that can usefully be attached to the antibodies of the present invention. 15 For flow cytometric applications, both for extracellular detection and for intracellular detection, common useful fluorophores can be fluorescein isothiocyanate (FITC), allophycocyanin (APC), R-phycoerythrin (PE), peridinin chlorophyll protein (PerCP), Texas Red, Cy3, Cy5, fluorescence resonance energy tandem fluorophores such as PerCP-Cy5.5, PE-Cy5, PE-Cy5.5, PE-Cy7, PE-Texas Red, and APC-Cy7.

20 Other fluorophores include, *inter alia*, Alexa Fluor® 350, Alexa Fluor® 488, Alexa Fluor® 532, Alexa Fluor® 546, Alexa Fluor® 568, Alexa Fluor® 594, Alexa Fluor® 647 (monoclonal antibody labeling kits available from Molecular Probes, Inc., Eugene, OR, USA), BODIPY dyes, such as BODIPY 493/503, BODIPY FL, BODIPY R6G, BODIPY 530/550, BODIPY TMR, BODIPY 558/568, BODIPY 558/568, BODIPY 25 564/570, BODIPY 576/589, BODIPY 581/591, BODIPY TR, BODIPY 630/650, BODIPY 650/665, Cascade Blue, Cascade Yellow, Dansyl, lissamine rhodamine B, Marina Blue, Oregon Green 488, Oregon Green 514, Pacific Blue, rhodamine 6G, rhodamine green, rhodamine red, tetramethylrhodamine, Texas Red (available from Molecular Probes, Inc., Eugene, OR, USA), and Cy2, Cy3, Cy3.5, Cy5, Cy5.5, Cy7, all of 30 which are also useful for fluorescently labeling the antibodies of the present invention. For secondary detection using labeled avidin, streptavidin, captavidin or neutravidin, the antibodies of the present invention can usefully be labeled with biotin.

When the antibodies of the present invention are used, e.g., for western blotting applications, they can usefully be labeled with radioisotopes, such as ^{33}P , ^{32}P , ^{35}S , ^3H , and ^{125}I . As another example, when the antibodies of the present invention are used for radioimmunotherapy, the label can usefully be ^{228}Th , ^{227}Ac , ^{225}Ac , ^{223}Ra , ^{213}Bi , ^{212}Pb ,
5 ^{212}Bi , ^{211}At , ^{203}Pb , ^{194}Os , ^{188}Re , ^{186}Re , ^{153}Sm , ^{149}Tb , ^{131}I , ^{125}I , ^{111}In , ^{105}Rh , $^{99\text{m}}\text{Tc}$, ^{97}Ru , ^{90}Y ,
 ^{90}Sr , ^{88}Y , ^{72}Se , ^{67}Cu , or ^{47}Sc .

As another example, when the antibodies of the present invention are to be used for *in vivo* diagnostic use, they can be rendered detectable by conjugation to MRI contrast agents, such as gadolinium diethylenetriaminepentaacetic acid (DTPA), Lauffer *et al.*,
10 *Radiology* 207(2): 529-38 (1998), or by radioisotopic labeling.

As would be understood, use of the labels described above is not restricted to the application as for which they were mentioned.

The antibodies of the present invention, including fragments and derivatives thereof, can also be conjugated to toxins, in order to target the toxin's ablative action to
15 cells that display and/or express the polypeptides of the present invention. Commonly, the antibody in such immunotoxins is conjugated to Pseudomonas exotoxin A, diphtheria toxin, shiga toxin A, anthrax toxin lethal factor, or ricin. See Hall (ed.), Immunotoxin Methods and Protocols (Methods in Molecular Biology, vol. 166), Humana Press (2000); and Frankel *et al.* (eds.), Clinical Applications of Immunotoxins, Springer-Verlag (1998).

20 The antibodies of the present invention can usefully be attached to a substrate, and it is, therefore, another aspect of the invention to provide antibodies that bind specifically to one or more of the polypeptides of the present invention, to one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, or the binding of which can be competitively inhibited by one or more of the polypeptides of
25 the present invention or one or more of the polypeptides encoded by the isolated nucleic acid molecules of the present invention, attached to a substrate. Substrates can be porous or nonporous, planar or nonplanar. For example, the antibodies of the present invention can usefully be conjugated to filtration media, such as NHS-activated Sepharose or CNBr-activated Sepharose for purposes of immunoaffinity chromatography. For example, the
30 antibodies of the present invention can usefully be attached to paramagnetic microspheres, typically by biotin-streptavidin interaction, which microsphere can then be used for isolation of cells that express or display the polypeptides of the present invention. As

another example, the antibodies of the present invention can usefully be attached to the surface of a microtiter plate for ELISA.

As noted above, the antibodies of the present invention can be produced in prokaryotic and eukaryotic cells. It is, therefore, another aspect of the present invention to provide cells that express the antibodies of the present invention, including hybridoma cells, B cells, plasma cells, and host cells recombinantly modified to express the antibodies of the present invention.

In yet a further aspect, the present invention provides aptamers evolved to bind specifically to one or more of the BSPs of the present invention or to polypeptides encoded by the BSNAs of the invention.

In sum, one of skill in the art, provided with the teachings of this invention, has available a variety of methods which may be used to alter the biological properties of the antibodies of this invention including methods which would increase or decrease the stability or half-life, immunogenicity, toxicity, affinity or yield of a given antibody molecule, or to alter it in any other way that may render it more suitable for a particular application.

Transgenic Animals and Cells

In another aspect, the invention provides transgenic cells and non-human organisms comprising nucleic acid molecules of the invention. In a preferred embodiment, the transgenic cells and non-human organisms comprise a nucleic acid molecule encoding a BSP. In a preferred embodiment, the BSP comprises an amino acid sequence selected from SEQ ID NO: 96-232, or a fragment, mutein, homologous protein or allelic variant thereof. In another preferred embodiment, the transgenic cells and non-human organism comprise a BSNA of the invention, preferably a BSNA comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-95, or a part, substantially similar nucleic acid molecule, allelic variant or hybridizing nucleic acid molecule thereof.

In another embodiment, the transgenic cells and non-human organisms have a targeted disruption or replacement of the endogenous orthologue of the human BSG. The transgenic cells can be embryonic stem cells or somatic cells. The transgenic non-human organisms can be chimeric, nonchimeric heterozygotes, and nonchimeric homozygotes. Methods of producing transgenic animals are well known in the art. *See, e.g., Hogan et*

al., Manipulating the Mouse Embryo: A Laboratory Manual, 2d ed., Cold Spring Harbor Press (1999); Jackson *et al.*, Mouse Genetics and Transgenics: A Practical Approach, Oxford University Press (2000); and Pinkert, Transgenic Animal Technology: A Laboratory Handbook, Academic Press (1999).

5 Any technique known in the art may be used to introduce a nucleic acid molecule of the invention into an animal to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection. (*see, e.g.*, Paterson *et al.*, *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver *et al.*, *Biotechnology* 11: 1263-1270 (1993); Wright *et al.*, *Biotechnology* 9: 830-834 (1991); and U.S. Patent No. 10 4,873,191, herein incorporated by reference in its entirety); retrovirus-mediated gene transfer into germ lines, blastocysts or embryos (*see, e.g.*, Van der Putten *et al.*, *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)); gene targeting in embryonic stem cells (*see, e.g.*, Thompson *et al.*, *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (*see, e.g.*, Lo, 1983, *Mol. Cell. Biol.* 3: 1803-1814 (1983)); introduction using a gene gun (*see, e.g.*, Ulmer *et al.*, *Science* 259: 1745-49 (1993); introducing nucleic acid constructs into 15 embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (*see, e.g.*, Lavitrano *et al.*, *Cell* 57: 717-723 (1989)).

Other techniques include, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (*see, e.g.*, 20 Campell *et al.*, *Nature* 380: 64-66 (1996); Wilmut *et al.*, *Nature* 385: 810-813 (1997)). The present invention provides for transgenic animals that carry the transgene (*i.e.*, a nucleic acid molecule of the invention) in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.* e., mosaic animals or chimeric animals.

The transgene may be integrated as a single transgene or as multiple copies, such 25 as in concatamers, *e. g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, *e.g.*, the teaching of Lasko *et al.* *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 6232- 6236 (1992). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant 30 gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the

transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (RT-PCR). Samples of transgenic gene-expressing tissue may also be evaluated

5 immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than

10 one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA

15 analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of

20 the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Methods for creating a transgenic animal with a disruption of a targeted gene are also well known in the art. In general, a vector is designed to comprise some nucleotide sequences homologous to the endogenous targeted gene. The vector is introduced into a cell so that it may integrate, via homologous recombination with chromosomal sequences, into the endogenous gene, thereby disrupting the function of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type. *See, e.g., Gu et al., Science 265: 103-106*

25 (1994). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. *See, e.g., Smithies et al., Nature 317: 230-234 (1985); Thomas et al., Cell 51: 503-*

30 *512 (1987); Thompson et al., Cell 5: 313-321 (1989).*

- In one embodiment, a mutant, non-functional nucleic acid molecule of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous nucleic acid sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable
- 5 marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications
- 10 to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene. *See, e.g., Thomas, supra* and *Thompson, supra*. However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.
- 15 In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.,* knockouts) are administered to a patient *in vivo*. Such cells may be obtained from an animal or patient or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells,
- 20 blood cells (*e.g.,* lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, *e.g.,* by transduction (using viral vectors, and preferably
- 25 vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve

30 expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, *e.g.,* in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, *e.g.*, genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. *See, e.g.*, U.S. Patent Nos. 5,399,349 and 5,460,959, each of which is 5 incorporated by reference herein in its entirety.

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of 10 components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with 15 aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Computer Readable Means

A further aspect of the invention is a computer readable means for storing the nucleic acid and amino acid sequences of the instant invention. In a preferred 20 embodiment, the invention provides a computer readable means for storing SEQ ID NO: 96-232 and SEQ ID NO: 1-95 as described herein, as the complete set of sequences or in any combination. The records of the computer readable means can be accessed for reading and display and for interface with a computer system for the application of programs allowing for the location of data upon a query for data meeting certain criteria, 25 the comparison of sequences, the alignment or ordering of sequences meeting a set of criteria, and the like.

The nucleic acid and amino acid sequences of the invention are particularly useful as components in databases useful for search analyses as well as in sequence analysis algorithms. As used herein, the terms "nucleic acid sequences of the invention" and 30 "amino acid sequences of the invention" mean any detectable chemical or physical characteristic of a polynucleotide or polypeptide of the invention that is or may be reduced to or stored in a computer readable form. These include, without limitation,

chromatographic scan data or peak data, photographic data or scan data therefrom, and mass spectrographic data.

This invention provides computer readable media having stored thereon sequences of the invention. A computer readable medium may comprise one or more of the following: a nucleic acid sequence comprising a sequence of a nucleic acid sequence of the invention; an amino acid sequence comprising an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of one or more nucleic acid sequences of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention; a set of nucleic acid sequences wherein at least one of said sequences comprises the sequence of a nucleic acid sequence of the invention; a set of amino acid sequences wherein at least one of said sequences comprises the sequence of an amino acid sequence of the invention; a data set representing a nucleic acid sequence comprising the sequence of a nucleic acid sequence of the invention; a data set representing a nucleic acid sequence encoding an amino acid sequence comprising the sequence of an amino acid sequence of the invention. The computer readable medium can be any composition of matter used to store information or data, including, for example, commercially available floppy disks, tapes, hard drives, compact disks, and video disks.

Also provided by the invention are methods for the analysis of character sequences, particularly genetic sequences. Preferred methods of sequence analysis include, for example, methods of sequence homology analysis, such as identity and similarity analysis, RNA structure analysis, sequence assembly, cladistic analysis, sequence motif analysis, open reading frame determination, nucleic acid base calling, and sequencing chromatogram peak analysis.

A computer-based method is provided for performing nucleic acid sequence identity or similarity identification. This method comprises the steps of providing a nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and comparing said nucleic acid sequence to at least one nucleic acid or amino acid sequence to identify sequence identity or similarity.

A computer-based method is also provided for performing amino acid homology identification, said method comprising the steps of: providing an amino acid sequence comprising the sequence of an amino acid of the invention in a computer readable medium; and comparing said amino acid sequence to at least one nucleic acid or an amino acid sequence to identify homology.

A computer-based method is still further provided for assembly of overlapping nucleic acid sequences into a single nucleic acid sequence, said method comprising the steps of: providing a first nucleic acid sequence comprising the sequence of a nucleic acid of the invention in a computer readable medium; and screening for at least one overlapping region between said first nucleic acid sequence and a second nucleic acid sequence. In addition, the invention includes a method of using patterns of expression associated with either the nucleic acids or proteins in a computer-based method to diagnose disease.

Diagnostic Methods for Breast Cancer

The present invention also relates to quantitative and qualitative diagnostic assays and methods for detecting, diagnosing, monitoring, staging and predicting cancers by comparing expression of a BSNA or a BSP in a human patient that has or may have breast cancer, or who is at risk of developing breast cancer, with the expression of a BSNA or a BSP in a normal human control. For purposes of the present invention, "expression of a BSNA" or "BSNA expression" means the quantity of BSNA mRNA that can be measured by any method known in the art or the level of transcription that can be measured by any method known in the art in a cell, tissue, organ or whole patient. Similarly, the term "expression of a BSP" or "BSP expression" means the amount of BSP that can be measured by any method known in the art or the level of translation of a BSNA that can be measured by any method known in the art.

The present invention provides methods for diagnosing breast cancer in a patient, by analyzing for changes in levels of BSNA or BSP in cells, tissues, organs or bodily fluids compared with levels of BSNA or BSP in cells, tissues, organs or bodily fluids of preferably the same type from a normal human control, wherein an increase, or decrease in certain cases, in levels of a BSNA or BSP in the patient versus the normal human control is associated with the presence of breast cancer or with a predilection to the disease. In another preferred embodiment, the present invention provides methods for diagnosing

breast cancer in a patient by analyzing changes in the structure of the mRNA of a BSG compared to the mRNA from a normal control. These changes include, without limitation, aberrant splicing, alterations in polyadenylation and/or alterations in 5' nucleotide capping. In yet another preferred embodiment, the present invention provides methods for 5 diagnosing breast cancer in a patient by analyzing changes in a BSP compared to a BSP from a normal patient. These changes include, e.g., alterations, including post translational modifications such as glycosylation and/or phosphorylation of the BSP or changes in the subcellular BSP localization.

For purposes of the present invention, diagnosing means that BSNA or BSP levels 10 are used to determine the presence or absence of disease in a patient. As will be understood by those of skill in the art, measurement of other diagnostic parameters may be required for definitive diagnosis or determination of the appropriate treatment for the disease. The determination may be made by a clinician, a doctor, a testing laboratory, or a patient using an over the counter test. The patient may have symptoms of disease or may 15 be asymptomatic. In addition, the BSNA or BSP levels of the present invention may be used as screening marker to determine whether further tests or biopsies are warranted. In addition, the BSNA or BSP levels may be used to determine the vulnerability or susceptibility to disease.

In a preferred embodiment, the expression of a BSNA is measured by determining 20 the amount of a mRNA that encodes an amino acid sequence selected from SEQ ID NO: 96-232, a homolog, an allelic variant, or a fragment thereof. In a more preferred embodiment, the BSNA expression that is measured is the level of expression of a BSNA mRNA selected from SEQ ID NO: 1-95, or a hybridizing nucleic acid, homologous nucleic acid or allelic variant thereof, or a part of any of these nucleic acid molecules. 25 BSNA expression may be measured by any method known in the art, such as those described *supra*, including measuring mRNA expression by Northern blot, quantitative or qualitative reverse transcriptase PCR (RT-PCR), microarray, dot or slot blots or *in situ* hybridization. See, e.g., Ausubel (1992), *supra*; Ausubel (1999), *supra*; Sambrook (1989), *supra*; and Sambrook (2001), *supra*. BSNA transcription may be measured by any 30 method known in the art including using a reporter gene hooked up to the promoter of a BSG of interest or doing nuclear run-off assays. Alterations in mRNA structure, e.g., aberrant splicing variants, may be determined by any method known in the art, including, RT-PCR followed by sequencing or restriction analysis. As necessary, BSNA expression

may be compared to a known control, such as normal breast nucleic acid, to detect a change in expression.

In another preferred embodiment, the expression of a BSP is measured by determining the level of a BSP having an amino acid sequence selected from the group 5 consisting of SEQ ID NO: 96-232, a homolog, an allelic variant, or a fragment thereof. Such levels are preferably determined in at least one of cells, tissues, organs and/or bodily fluids, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for diagnosing over- or underexpression of a BSNA or BSP compared to normal control bodily fluids, cells, or tissue samples may 10 be used to diagnose the presence of breast cancer. The expression level of a BSP may be determined by any method known in the art, such as those described *supra*. In a preferred embodiment, the BSP expression level may be determined by radioimmunoassays, competitive-binding assays, ELISA, Western blot, FACS, immunohistochemistry, immunoprecipitation, proteomic approaches: two-dimensional gel electrophoresis (2D 15 electrophoresis) and non-gel-based approaches such as mass spectrometry or protein interaction profiling. See, e.g., Harlow (1999), *supra*; Ausubel (1992), *supra*; and Ausubel (1999), *supra*. Alterations in the BSP structure may be determined by any method known in the art, including, e.g., using antibodies that specifically recognize phosphoserine, phosphothreonine or phosphotyrosine residues, two-dimensional polyacrylamide gel 20 electrophoresis (2D PAGE) and/or chemical analysis of amino acid residues of the protein.

Id.

In a preferred embodiment, a radioimmunoassay (RIA) or an ELISA is used. An antibody specific to a BSP is prepared if one is not already available. In a preferred embodiment, the antibody is a monoclonal antibody. The anti-BSP antibody is bound to a 25 solid support and any free protein binding sites on the solid support are blocked with a protein such as bovine serum albumin. A sample of interest is incubated with the antibody on the solid support under conditions in which the BSP will bind to the anti-BSP antibody. The sample is removed, the solid support is washed to remove unbound material, and an anti-BSP antibody that is linked to a detectable reagent (a radioactive substance for RIA 30 and an enzyme for ELISA) is added to the solid support and incubated under conditions in which binding of the BSP to the labeled antibody will occur. After binding, the unbound labeled antibody is removed by washing. For an ELISA, one or more substrates are added to produce a colored reaction product that is based upon the amount of a BSP in the

sample. For an RIA, the solid support is counted for radioactive decay signals by any method known in the art. Quantitative results for both RIA and ELISA typically are obtained by reference to a standard curve.

Other methods to measure BSP levels are known in the art. For instance, a 5 competition assay may be employed wherein an anti-BSP antibody is attached to a solid support and an allocated amount of a labeled BSP and a sample of interest are incubated with the solid support. The amount of labeled BSP attached to the solid support can be correlated to the quantity of a BSP in the sample.

Of the proteomic approaches, 2D PAGE is a well known technique. Isolation of 10 individual proteins from a sample such as serum is accomplished using sequential separation of proteins by isoelectric point and molecular weight. Typically, polypeptides are first separated by isoelectric point (the first dimension) and then separated by size using an electric current (the second dimension). In general, the second dimension is perpendicular to the first dimension. Because no two proteins with different sequences are 15 identical on the basis of both size and charge, the result of 2D PAGE is a roughly square gel in which each protein occupies a unique spot. Analysis of the spots with chemical or antibody probes, or subsequent protein microsequencing can reveal the relative abundance of a given protein and the identity of the proteins in the sample.

Expression levels of a BSNA can be determined by any method known in the art, 20 including PCR and other nucleic acid methods, such as ligase chain reaction (LCR) and nucleic acid sequence based amplification (NASBA), can be used to detect malignant cells for diagnosis and monitoring of various malignancies. For example, reverse-transcriptase PCR (RT-PCR) is a powerful technique which can be used to detect the presence of a specific mRNA population in a complex mixture of thousands of other mRNA species. In 25 RT-PCR, an mRNA species is first reverse transcribed to complementary DNA (cDNA) with use of the enzyme reverse transcriptase; the cDNA is then amplified as in a standard PCR reaction.

Hybridization to specific DNA molecules (*e.g.*, oligonucleotides) arrayed on a solid support can be used to both detect the expression of and quantitate the level of 30 expression of one or more BSNA's of interest. In this approach, all or a portion of one or more BSNA's is fixed to a substrate. A sample of interest, which may comprise RNA, *e.g.*, total RNA or polyA-selected mRNA, or a complementary DNA (cDNA) copy of the RNA is incubated with the solid support under conditions in which hybridization will occur

between the DNA on the solid support and the nucleic acid molecules in the sample of interest. Hybridization between the substrate-bound DNA and the nucleic acid molecules in the sample can be detected and quantitated by several means, including, without limitation, radioactive labeling or fluorescent labeling of the nucleic acid molecule or a secondary molecule designed to detect the hybrid.

The above tests can be carried out on samples derived from a variety of cells, bodily fluids and/or tissue extracts such as homogenates or solubilized tissue obtained from a patient. Tissue extracts are obtained routinely from tissue biopsy and autopsy material. Bodily fluids useful in the present invention include blood, urine, saliva or any other bodily secretion or derivative thereof. As used herein "blood" includes whole blood, plasma, serum, circulating epithelial cells, constituents, or any derivative of blood.

In addition to detection in bodily fluids, the proteins and nucleic acids of the invention are suitable to detection by cell capture technology. Whole cells may be captured by a variety methods for example magnetic separation, such as described in U.S. Patent. Nos. 5,200,084; 5,186,827; 5,108,933; and 4,925,788, the disclosures of which are incorporated herein by reference in their entireties. Epithelial cells may be captured using such products as Dynabeads® or CELLlection™ (Dynal Biotech, Oslo, Norway). Alternatively, fractions of blood may be captured, e.g., the buffy coat fraction (50mm cells isolated from 5ml of blood) containing epithelial cells. In addition, cancer cells may be captured using the techniques described in WO 00/47998, the disclosure of which is incorporated herein by reference in its entirety. Once the cells are captured or concentrated, the proteins or nucleic acids are detected by the means described in the subject application. Alternatively, nucleic acids may be captured directly from blood samples, see U.S. Patent Nos. 6,156,504, 5,501,963; or WO 01/42504, the disclosures of which are incorporated herein by reference in their entireties.

In a preferred embodiment, the specimen tested for expression of BSNA or BSP includes without limitation breast tissue, breast cells grown in cell culture, blood, serum, lymph node tissue, and lymphatic fluid. In another preferred embodiment, especially when metastasis of a primary breast cancer is known or suspected, specimens include, without limitation, tissues from brain, bone, bone marrow, liver, lungs, colon, and adrenal glands. In general, the tissues may be sampled by biopsy, including, without limitation, needle biopsy, e.g., transthoracic needle aspiration, cervical mediastinoscopy, endoscopic

lymph node biopsy, video-assisted thoracoscopy, exploratory thoracotomy, bone marrow biopsy and bone marrow aspiration.

All the methods of the present invention may optionally include determining the expression levels of one or more other cancer markers in addition to determining the expression level of a BSNA or BSP. In many cases, the use of another cancer marker will decrease the likelihood of false positives or false negatives. In one embodiment, the one or more other cancer markers include other BSNA or BSPs as disclosed herein. Other cancer markers useful in the present invention will depend on the cancer being tested and are known to those of skill in the art. In a preferred embodiment, at least one other cancer marker in addition to a particular BSNA or BSP is measured. In a more preferred embodiment, at least two other additional cancer markers are used. In an even more preferred embodiment, at least three, more preferably at least five, even more preferably at least ten additional cancer markers are used.

Diagnosing

In one aspect, the invention provides a method for determining the expression levels and/or structural alterations of one or more BSNA and/or BSP in a sample from a patient suspected of having breast cancer. In general, the method comprises the steps of obtaining the sample from the patient, determining the expression level or structural alterations of a BSNA and/or BSP and then ascertaining whether the patient has breast cancer from the expression level of the BSNA or BSP. In general, if high expression relative to a control of a BSNA or BSP is indicative of breast cancer, a diagnostic assay is considered positive if the level of expression of the BSNA or BSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a BSNA or BSP is indicative of breast cancer, a diagnostic assay is considered positive if the level of expression of the BSNA or BSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

The present invention also provides a method of determining whether breast cancer has metastasized in a patient. One may identify whether the breast cancer has metastasized by measuring the expression levels and/or structural alterations of one or more BSNA and/or BSPs in a variety of tissues. The presence of a BSNA or BSP in a tissue other than breast at levels higher than that of corresponding noncancerous tissue (e.g., the same tissue from another individual) is indicative of metastasis if high level expression of a BSNA or BSP is associated with breast cancer. Similarly, the presence of a BSNA or BSP in a tissue other than breast at levels lower than that of corresponding noncancerous tissue is indicative of metastasis if low level expression of a BSNA or BSP is associated with breast cancer. Further, the presence of a structurally altered BSNA or BSP that is associated with breast cancer is also indicative of metastasis.

In general, if high expression relative to a control of a BSNA or BSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the BSNA or BSP is at least one and a half times higher, and more preferably are at least two times higher, still more preferably five times higher, even more preferably at least ten times higher, than in preferably the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a BSNA or BSP is indicative of metastasis, an assay for metastasis is considered positive if the level of expression of the BSNA or BSP is at least one and a half times lower, and more preferably are at least two times lower, still more preferably five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid of a normal human control.

Staging

The invention also provides a method of staging breast cancer in a human patient. The method comprises identifying a human patient having breast cancer and analyzing cells, tissues or bodily fluids from such human patient for expression levels and/or structural alterations of one or more BSNA and/or BSPs. First, one or more tumors from a variety of patients are staged according to procedures well known in the art, and the expression levels of one or more BSNA and/or BSPs is determined for each stage to obtain a standard expression level for each BSNA and BSP. Then, the BSNA or BSP expression levels of the BSNA or BSP are determined in a biological sample from a patient whose stage of cancer is not known. The BSNA or BSP expression levels from the patient are

then compared to the standard expression level. By comparing the expression level of the BSNA and BSPs from the patient to the standard expression levels, one may determine the stage of the tumor. The same procedure may be followed using structural alterations of a BSNA or BSP to determine the stage of a breast cancer.

5 *Monitoring*

Further provided is a method of monitoring breast cancer in a human patient. One may monitor a human patient to determine whether there has been metastasis and, if there has been, when metastasis began to occur. One may also monitor a human patient to determine whether a preneoplastic lesion has become cancerous. One may also monitor a
10 human patient to determine whether a therapy, e.g., chemotherapy, radiotherapy or surgery, has decreased or eliminated the breast cancer. The monitoring may determine if there has been a reoccurrence and, if so, determine its nature. The method comprises identifying a human patient that one wants to monitor for breast cancer, periodically analyzing cells, tissues or bodily fluids from such human patient for expression levels of
15 one or more BSNA or BSPs, and comparing the BSNA or BSP levels over time to those BSNA or BSP expression levels obtained previously. Patients may also be monitored by measuring one or more structural alterations in a BSNA or BSP that are associated with breast cancer.

If increased expression of a BSNA or BSP is associated with metastasis, treatment
20 failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting an increase in the expression level of a BSNA or BSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. One having ordinary skill in the art would recognize that if this were the case, then a decreased expression level would be indicative of no metastasis, effective therapy or failure to
25 progress to a neoplastic lesion. If decreased expression of a BSNA or BSP is associated with metastasis, treatment failure, or conversion of a preneoplastic lesion to a cancerous lesion, then detecting a decrease in the expression level of a BSNA or BSP indicates that the tumor is metastasizing, that treatment has failed or that the lesion is cancerous, respectively. In a preferred embodiment, the levels of BSNA or BSPs are determined
30 from the same cell type, tissue or bodily fluid as prior patient samples. Monitoring a patient for onset of breast cancer metastasis is periodic and preferably is done on a quarterly basis, but may be done more or less frequently.

The methods described herein can further be utilized as prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with increased or decreased expression levels of a BSNA and/or BSP. The present invention provides a method in which a test sample is obtained from a human patient and one or 5 more BSNA and/or BSPs are detected. The presence of higher (or lower) BSNA or BSP levels as compared to normal human controls is diagnostic for the human patient being at risk for developing cancer, particularly breast cancer. The effectiveness of therapeutic agents to decrease (or increase) expression or activity of one or more BSNA and/or BSPs of the invention can also be monitored by analyzing levels of expression of the BSNA and/or BSPs in a human patient in clinical trials or in *in vitro* screening assays such as in 10 human cells. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the human patient or cells, as the case may be, to the agent being tested.

Detection of Genetic Lesions or Mutations

15 The methods of the present invention can also be used to detect genetic lesions or mutations in a BSG, thereby determining if a human with the genetic lesion is susceptible to developing breast cancer or to determine what genetic lesions are responsible, or are partly responsible, for a person's existing breast cancer. Genetic lesions can be detected, for example, by ascertaining the existence of a deletion, insertion and/or substitution of 20 one or more nucleotides from the BSGs of this invention, a chromosomal rearrangement of a BSG, an aberrant modification of a BSG (such as of the methylation pattern of the genomic DNA), or allelic loss of a BSG. Methods to detect such lesions in the BSG of this invention are known to those having ordinary skill in the art following the teachings of the specification.

25 **Methods of Detecting Noncancerous Breast Diseases**

The present invention also provides methods for determining the expression levels and/or structural alterations of one or more BSNA and/or BSPs in a sample from a patient suspected of having or known to have a noncancerous breast disease. In general, the method comprises the steps of obtaining a sample from the patient, determining the 30 expression level or structural alterations of a BSNA and/or BSP, comparing the expression level or structural alteration of the BSNA or BSP to a normal breast control, and then ascertaining whether the patient has a noncancerous breast disease. In general, if high

expression relative to a control of a BSNA or BSP is indicative of a particular noncancerous breast disease, a diagnostic assay is considered positive if the level of expression of the BSNA or BSP is at least two times higher, and more preferably are at least five times higher, even more preferably at least ten times higher, than in preferably 5 the same cells, tissues or bodily fluid of a normal human control. In contrast, if low expression relative to a control of a BSNA or BSP is indicative of a noncancerous breast disease, a diagnostic assay is considered positive if the level of expression of the BSNA or BSP is at least two times lower, more preferably are at least five times lower, even more preferably at least ten times lower than in preferably the same cells, tissues or bodily fluid 10 of a normal human control. The normal human control may be from a different patient or from uninvolved tissue of the same patient.

One having ordinary skill in the art may determine whether a BSNA and/or BSP is associated with a particular noncancerous breast disease by obtaining breast tissue from a patient having a noncancerous breast disease of interest and determining which BSNAs and/or BSPs are expressed in the tissue at either a higher or a lower level than in normal 15 breast tissue. In another embodiment, one may determine whether a BSNA or BSP exhibits structural alterations in a particular noncancerous breast disease state by obtaining breast tissue from a patient having a noncancerous breast disease of interest and determining the structural alterations in one or more BSNAs and/or BSPs relative to 20 normal breast tissue.

Methods for Identifying Breast Tissue

In another aspect, the invention provides methods for identifying breast tissue. These methods are particularly useful in, *e.g.*, forensic science, breast cell differentiation and development, and in tissue engineering.

25 In one embodiment, the invention provides a method for determining whether a sample is breast tissue or has breast tissue-like characteristics. The method comprises the steps of providing a sample suspected of comprising breast tissue or having breast tissue-like characteristics, determining whether the sample expresses one or more BSNAs and/or BSPs, and, if the sample expresses one or more BSNAs and/or BSPs, concluding that the 30 sample comprises breast tissue. In a preferred embodiment, the BSNA encodes a polypeptide having an amino acid sequence selected from SEQ ID NO: 96-232, or a homolog, allelic variant or fragment thereof. In a more preferred embodiment, the BSNA

has a nucleotide sequence selected from SEQ ID NO: 1-95, or a hybridizing nucleic acid, an allelic variant or a part thereof. Determining whether a sample expresses a BSNA can be accomplished by any method known in the art. Preferred methods include hybridization to microarrays, Northern blot hybridization, and quantitative or qualitative 5 RT-PCR. In another preferred embodiment, the method can be practiced by determining whether a BSP is expressed. Determining whether a sample expresses a BSP can be accomplished by any method known in the art. Preferred methods include Western blot, ELISA, RIA and 2D PAGE. In one embodiment, the BSP has an amino acid sequence selected from SEQ ID NO: 96-232, or a homolog, allelic variant or fragment thereof. In 10 another preferred embodiment, the expression of at least two BSNAs and/or BSPs is determined. In a more preferred embodiment, the expression of at least three, more preferably four and even more preferably five BSNAs and/or BSPs are determined.

In one embodiment, the method can be used to determine whether an unknown tissue is breast tissue. This is particularly useful in forensic science, in which small, 15 damaged pieces of tissues that are not identifiable by microscopic or other means are recovered from a crime or accident scene. In another embodiment, the method can be used to determine whether a tissue is differentiating or developing into breast tissue. This is important in monitoring the effects of the addition of various agents to cell or tissue culture, *e.g.*, in producing new breast tissue by tissue engineering. These agents include, 20 *e.g.*, growth and differentiation factors, extracellular matrix proteins and culture medium. Other factors that may be measured for effects on tissue development and differentiation include gene transfer into the cells or tissues, alterations in pH, aqueous:air interface and various other culture conditions.

Methods for Producing and Modifying Breast Tissue

25 In another aspect, the invention provides methods for producing engineered breast tissue or cells. In one embodiment, the method comprises the steps of providing cells, introducing a BSNA or a BSG into the cells, and growing the cells under conditions in which they exhibit one or more properties of breast tissue cells. In a preferred embodiment, the cells are pluripotent. As is well known in the art, normal breast tissue 30 comprises a large number of different cell types. Thus, in one embodiment, the engineered breast tissue or cells comprises one of these cell types. In another embodiment, the engineered breast tissue or cells comprises more than one breast cell

type. Further, the culture conditions of the cells or tissue may require manipulation in order to achieve full differentiation and development of the breast cell tissue. Methods for manipulating culture conditions are well known in the art.

Nucleic acid molecules encoding one or more BSPs are introduced into cells, 5 preferably pluripotent cells. In a preferred embodiment, the nucleic acid molecules encode BSPs having amino acid sequences selected from SEQ ID NO: 96-232, or homologous proteins, analogs, allelic variants or fragments thereof. In a more preferred embodiment, the nucleic acid molecules have a nucleotide sequence selected from SEQ ID NO: 1-95, or hybridizing nucleic acids, allelic variants or parts thereof. In another highly 10 preferred embodiment, a BSG is introduced into the cells. Expression vectors and methods of introducing nucleic acid molecules into cells are well known in the art and are described in detail, *supra*.

Artificial breast tissue may be used to treat patients who have lost some or all of their breast function.

15 **Pharmaceutical Compositions**

In another aspect, the invention provides pharmaceutical compositions comprising the nucleic acid molecules, polypeptides, fusion proteins, antibodies, antibody derivatives, antibody fragments, agonists, antagonists, or inhibitors of the present invention. In a preferred embodiment, the pharmaceutical composition comprises a BSNA or part thereof. 20 In a more preferred embodiment, the BSNA has a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-95, a nucleic acid that hybridizes thereto, an allelic variant thereof, or a nucleic acid that has substantial sequence identity thereto. In another preferred embodiment, the pharmaceutical composition comprises a BSP or fragment thereof. In a more preferred embodiment, the pharmaceutical composition comprises a 25 BSP having an amino acid sequence that is selected from the group consisting of SEQ ID NO: 96-232, a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof. In another preferred embodiment, the pharmaceutical composition comprises an anti-BSP antibody, preferably an antibody that specifically binds to a BSP having an amino acid that is selected from the 30 group consisting of SEQ ID NO: 96-232, or an antibody that binds to a polypeptide that is homologous thereto, a fusion protein comprising all or a portion of the polypeptide, or an analog or derivative thereof.

Due to the association of angiogenesis with cancer vascularization there is great need of new markers and methods for diagnosing angiogenesis activity to identify developing tumors and angiogenesis related diseases. Furthermore, great need is also present for new molecular targets useful in the treatment of angiogenesis and angiogenesis

- 5 related diseases such as cancer. In addition known modulators of angiogenesis such as endostatin or vascular endothelial growth factor (VEGF). Use of the methods and compositions disclosed herein in combination with anti-angiogenesis drugs, drugs that block the matrix breakdown (such as BMS-275291, Dalteparin (Fragmin®), Suramin), drugs that inhibit endothelial cells (2-methoxyestradiol (2-ME), CC-5013 (Thalidomide 10 Analog), Combretastatin A4 Phosphate, LY317615 (Protein Kinase C Beta Inhibitor), Soy Isoflavone (Genistein; Soy Protein Isolate), Thalidomide), drugs that block activators of angiogenesis (AE-941 (Neovastat™; GW786034), Anti-VEGF Antibody (Bevacizumab; Avastin™), Interferon-alpha, PTK787/ZK 222584, VEGF-Trap, ZD6474), Drugs that inhibit endothelial-specific integrin/survival signaling (EMD 121974, Anti-Anb3 Integrin 15 Antibody (Medi-522; Vitaxin™)).

Such a composition typically contains from about 0.1 to 90% by weight of a therapeutic agent of the invention formulated in and/or with a pharmaceutically acceptable carrier or excipient.

- Pharmaceutical formulation is a well-established art that is further described in 20 Gennaro (ed.), Remington: The Science and Practice of Pharmacy, 20th ed., Lippincott, Williams & Wilkins (2000); Ansel *et al.*, Pharmaceutical Dosage Forms and Drug Delivery Systems, 7th ed., Lippincott Williams & Wilkins (1999); and Kibbe (ed.), Handbook of Pharmaceutical Excipients American Pharmaceutical Association, 3rd ed. (2000) and thus need not be described in detail herein.

- Briefly, formulation of the pharmaceutical compositions of the present invention 25 will depend upon the route chosen for administration. The pharmaceutical compositions utilized in this invention can be administered by various routes including both enteral and parenteral routes, including oral, intravenous, intramuscular, subcutaneous, inhalation, topical, sublingual, rectal, intra-arterial, intramedullary, intrathecal, intraventricular, 30 transmucosal, transdermal, intranasal, intraperitoneal, intrapulmonary, and intrauterine.

Oral dosage forms can be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Solid formulations of the compositions for oral administration can contain suitable carriers or excipients, such as carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, or microcrystalline cellulose; gums including arabic and tragacanth; proteins such as gelatin and collagen; inorganics, such as kaolin, calcium carbonate, dicalcium phosphate, sodium chloride; and other agents such as acacia and alginic acid.

Agents that facilitate disintegration and/or solubilization can be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate, microcrystalline cellulose, cornstarch, sodium starch glycolate, and alginic acid.

Tablet binders that can be used include acacia, methylcellulose, sodium carboxymethylcellulose, polyvinylpyrrolidone (PovidoneTM), hydroxypropyl methylcellulose, sucrose, starch and ethylcellulose.

Lubricants that can be used include magnesium stearates, stearic acid, silicone fluid, talc, waxes, oils, and colloidal silica.

Fillers, agents that facilitate disintegration and/or solubilization, tablet binders and lubricants, including the aforementioned, can be used singly or in combination.

Solid oral dosage forms need not be uniform throughout. For example, dragee cores can be used in conjunction with suitable coatings, such as concentrated sugar solutions, which can also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures.

Oral dosage forms of the present invention include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds can be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Additionally, dyestuffs or pigments can be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, *i.e.*, dosage.

Liquid formulations of the pharmaceutical compositions for oral (enteral) administration are prepared in water or other aqueous vehicles and can contain various suspending agents such as methylcellulose, alginates, tragacanth, pectin, kelgin, carrageenan, acacia, polyvinylpyrrolidone, and polyvinyl alcohol. The liquid formulations can also include solutions, emulsions, syrups and elixirs containing, together with the active compound(s), wetting agents, sweeteners, and coloring and flavoring agents.

The pharmaceutical compositions of the present invention can also be formulated for parenteral administration. Formulations for parenteral administration can be in the form of aqueous or non-aqueous isotonic sterile injection solutions or suspensions.

For intravenous injection, water soluble versions of the compounds of the present invention are formulated in, or if provided as a lyophilate, mixed with, a physiologically acceptable fluid vehicle, such as 5% dextrose ("D5"), physiologically buffered saline, 0.9% saline, Hanks' solution, or Ringer's solution. Intravenous formulations may include carriers, excipients or stabilizers including, without limitation, calcium, human serum albumin, citrate, acetate, calcium chloride, carbonate, and other salts.

Intramuscular preparations, e.g. a sterile formulation of a suitable soluble salt form of the compounds of the present invention, can be dissolved and administered in a pharmaceutical excipient such as Water-for-Injection, 0.9% saline, or 5% glucose solution. Alternatively, a suitable insoluble form of the compound can be prepared and administered as a suspension in an aqueous base or a pharmaceutically acceptable oil base, such as an ester of a long chain fatty acid (e.g., ethyl oleate), fatty oils such as sesame oil, triglycerides, or liposomes.

Parenteral formulations of the compositions can contain various carriers such as vegetable oils, dimethylacetamide, dimethylformamide, ethyl lactate, ethyl carbonate, isopropyl myristate, ethanol, polyols (glycerol, propylene glycol, liquid polyethylene glycol, and the like).

Aqueous injection suspensions can also contain substances that increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Non-lipid polycationic amino polymers can also be used for delivery. Optionally, the suspension can also contain suitable stabilizers or agents that increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical compositions of the present invention can also be formulated to permit injectable, long-term, deposition. Injectable depot forms may be made by forming

microencapsulated matrices of the compound in biodegradable polymers such as polylactide-polyglycolide. Depending upon the ratio of drug to polymer and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot 5 injectable formulations are also prepared by entrapping the drug in microemulsions that are compatible with body tissues.

The pharmaceutical compositions of the present invention can be administered topically. For topical use the compounds of the present invention can also be prepared in suitable forms to be applied to the skin, or mucus membranes of the nose and throat, and 10 can take the form of lotions, creams, ointments, liquid sprays or inhalants, drops, tinctures, lozenges, or throat paints. Such topical formulations further can include chemical compounds such as dimethylsulfoxide (DMSO) to facilitate surface penetration of the active ingredient. In other transdermal formulations, typically in patch-delivered 15 formulations, the pharmaceutically active compound is formulated with one or more skin penetrants, such as 2-N-methyl-pyrrolidone (NMP) or Azone. A topical semi-solid ointment formulation typically contains a concentration of the active ingredient from about 1 to 20%, e.g., 5 to 10%, in a carrier such as a pharmaceutical cream base.

For application to the eyes or ears, the compounds of the present invention can be presented in liquid or semi-liquid form formulated in hydrophobic or hydrophilic bases as 20 ointments, creams, lotions, paints or powders.

For rectal administration the compounds of the present invention can be administered in the form of suppositories admixed with conventional carriers such as cocoa butter, wax or other glyceride.

Inhalation formulations can also readily be formulated. For inhalation, various 25 powder and liquid formulations can be prepared. For aerosol preparations, a sterile formulation of the compound or salt form of the compound may be used in inhalers, such as metered dose inhalers, and nebulizers. Aerosolized forms may be especially useful for treating respiratory disorders.

Alternatively, the compounds of the present invention can be in powder form for 30 reconstitution in the appropriate pharmaceutically acceptable carrier at the time of delivery.

The pharmaceutically active compound in the pharmaceutical compositions of the present invention can be provided as the salt of a variety of acids, including but not limited

to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms.

5 After pharmaceutical compositions have been prepared, they are packaged in an appropriate container and labeled for treatment of an indicated condition.

The active compound will be present in an amount effective to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

A "therapeutically effective dose" refers to that amount of active ingredient, for 10 example BSP polypeptide, fusion protein, or fragments thereof, antibodies specific for BSP, agonists, antagonists or inhibitors of BSP, which ameliorates the signs or symptoms of the disease or prevent progression thereof; as would be understood in the medical arts, cure, although desired, is not required.

15 The therapeutically effective dose of the pharmaceutical agents of the present invention can be estimated initially by *in vitro* tests, such as cell culture assays, followed by assay in model animals, usually mice, rats, rabbits, dogs, or pigs. The animal model can also be used to determine an initial preferred concentration range and route of administration.

20 For example, the ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population) can be determined in one or more cell culture of animal model systems. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as LD50/ED50. Pharmaceutical compositions that exhibit large therapeutic indices are preferred.

25 The data obtained from cell culture assays and animal studies are used in formulating an initial dosage range for human use, and preferably provide a range of circulating concentrations that includes the ED50 with little or no toxicity. After administration, or between successive administrations, the circulating concentration of active agent varies within this range depending upon pharmacokinetic factors well known in the art, such as the dosage form employed, sensitivity of the patient, and the route of 30 administration.

The exact dosage will be determined by the practitioner, in light of factors specific to the subject requiring treatment. Factors that can be taken into account by the practitioner include the severity of the disease state, general health of the subject, age,

weight, gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions can be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

5 Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Where the therapeutic agent is a protein or antibody of the present invention, the therapeutic protein or antibody agent typically is administered at a daily dosage of 0.01 mg to 30 mg/kg of body weight of the patient (e.g., 1mg/kg to 5 mg/kg). The pharmaceutical formulation can be
10 administered in multiple doses per day, if desired, to achieve the total desired daily dose.

Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells,
15 conditions, locations, etc.

Conventional methods, known to those of ordinary skill in the art of medicine, can be used to administer the pharmaceutical formulation(s) of the present invention to the patient. The pharmaceutical compositions of the present invention can be administered alone, or in combination with other therapeutic agents or interventions.

20 **Therapeutic Methods**

The present invention further provides methods of treating subjects having defects in a gene of the invention, e.g., in expression, activity, distribution, localization, and/or solubility, which can manifest as a disorder of breast function. As used herein, "treating" includes all medically-acceptable types of therapeutic intervention, including palliation
25 and prophylaxis (prevention) of disease. The term "treating" encompasses any improvement of a disease, including minor improvements. These methods are discussed below.

Gene Therapy and Vaccines

The isolated nucleic acids of the present invention can also be used to drive *in vivo*
30 expression of the polypeptides of the present invention. *In vivo* expression can be driven from a vector, typically a viral vector, often a vector based upon a replication incompetent retrovirus, an adenovirus, or an adeno-associated virus (AAV), for the purpose of gene

therapy. *In vivo* expression can also be driven from signals endogenous to the nucleic acid or from a vector, often a plasmid vector, such as pVAX1 (Invitrogen, Carlsbad, CA, USA), for purpose of “naked” nucleic acid vaccination, as further described in U.S. Patent Nos. 5,589,466; 5,679,647; 5,804,566; 5,830,877; 5,843,913; 5,880,104; 5,958,891; 5,985,847; 6,017,897; 6,110,898; 6,204,250, the disclosures of which are incorporated herein by reference in their entireties. For cancer therapy, it is preferred that the vector also be tumor-selective. *See, e.g.*, Doronin *et al.*, *J. Virol.* 75: 3314-24 (2001).

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a nucleic acid molecule of the present invention is administered. The nucleic acid molecule can be delivered in a vector that drives expression of a BSP, fusion protein, or fragment thereof, or without such vector. Nucleic acid compositions that can drive expression of a BSP are administered, for example, to complement a deficiency in the native BSP, or as DNA vaccines. Expression vectors derived from virus, replication deficient retroviruses, adenovirus, adeno-associated (AAV) virus, herpes virus, or vaccinia virus can be used as can plasmids. *See, e.g.*, Cid-Arregui, *supra*. In a preferred embodiment, the nucleic acid molecule encodes a BSP having the amino acid sequence of SEQ ID NO: 96-232, or a fragment, fusion protein, allelic variant or homolog thereof.

In still other therapeutic methods of the present invention, pharmaceutical compositions comprising host cells that express a BSP, fusions, or fragments thereof can be administered. In such cases, the cells are typically autologous, so as to circumvent xenogeneic or allotypic rejection, and are administered to complement defects in BSP production or activity. In a preferred embodiment, the nucleic acid molecules in the cells encode a BSP having the amino acid sequence of SEQ ID NO: 96-232, or a fragment, fusion protein, allelic variant or homolog thereof.

Antisense Administration

Antisense nucleic acid compositions, or vectors that drive expression of a BSG antisense nucleic acid, are administered to downregulate transcription and/or translation of a BSG in circumstances in which excessive production, or production of aberrant protein, is the pathophysiologic basis of disease.

Antisense compositions useful in therapy can have a sequence that is complementary to coding or to noncoding regions of a BSG. For example,

oligonucleotides derived from the transcription initiation site, *e.g.*, between positions -10 and +10 from the start site, are preferred.

- Catalytic antisense compositions, such as ribozymes, that are capable of sequence-specific hybridization to BSG transcripts, are also useful in therapy. *See, e.g.*,
- 5 Phylactou, *Adv. Drug Deliv. Rev.* 44(2-3): 97-108 (2000); Phylactou *et al.*, *Hum. Mol. Genet.* 7(10): 1649-53 (1998); Rossi, *Ciba Found. Symp.* 209: 195-204 (1997); and Sigurdsson *et al.*, *Trends Biotechnol.* 13(8): 286-9 (1995).

Other nucleic acids useful in the therapeutic methods of the present invention are those that are capable of triplex helix formation in or near the BSG genomic locus. Such 10 triplexing oligonucleotides are able to inhibit transcription. *See, e.g.*, Intody *et al.*, *Nucleic Acids Res.* 28(21): 4283-90 (2000); and McGuffie *et al.*, *Cancer Res.* 60(14): 3790-9 (2000). Pharmaceutical compositions comprising such triplex forming oligos (TFOs) are administered in circumstances in which excessive production, or production of aberrant protein, is a pathophysiologic basis of disease.

15 In a preferred embodiment, the antisense molecule is derived from a nucleic acid molecule encoding a BSP, preferably a BSP comprising an amino acid sequence of SEQ ID NO: 96-232, or a fragment, allelic variant or homolog thereof. In a more preferred embodiment, the antisense molecule is derived from a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or 20 hybridizing nucleic acid thereof.

Polypeptide Administration

In one embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising a BSP, a fusion protein, fragment, analog or derivative thereof is administered to a subject with a 25 clinically-significant BSP defect.

Protein compositions are administered, for example, to complement a deficiency in native BSP. In other embodiments, protein compositions are administered as a vaccine to elicit a humoral and/or cellular immune response to BSP. The immune response can be used to modulate activity of BSP or, depending on the immunogen, to immunize against 30 aberrant or aberrantly expressed forms, such as mutant or inappropriately expressed isoforms. In yet other embodiments, protein fusions having a toxic moiety are administered to ablate cells that aberrantly accumulate BSP.

In a preferred embodiment, the polypeptide administered is a BSP comprising an amino acid sequence of SEQ ID NO: 96-232, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the polypeptide is encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, 5 allelic variant, substantially similar or hybridizing nucleic acid thereof.

Antibody, Agonist and Antagonist Administration

In another embodiment of the therapeutic methods of the present invention, a therapeutically effective amount of a pharmaceutical composition comprising an antibody (including fragment or derivative thereof) of the present invention is administered. As is 10 well known, antibody compositions are administered, for example, to antagonize activity of BSP, or to target therapeutic agents to sites of BSP presence and/or accumulation. In a preferred embodiment, the antibody specifically binds to a BSP comprising an amino acid sequence of SEQ ID NO: 96-232, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the antibody specifically binds to a 15 BSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

The present invention also provides methods for identifying modulators which bind to a BSP or have a modulatory effect on the expression or activity of a BSP. Modulators which decrease the expression or activity of BSP (antagonists) are believed to 20 be useful in treating breast cancer. Such screening assays are known to those of skill in the art and include, without limitation, cell-based assays and cell-free assays. Small molecules predicted via computer imaging to specifically bind to regions of a BSP can also be designed, synthesized and tested for use in the imaging and treatment of breast cancer. Further, libraries of molecules can be screened for potential anticancer agents by 25 assessing the ability of the molecule to bind to the BSPs identified herein. Molecules identified in the library as being capable of binding to a BSP are key candidates for further evaluation for use in the treatment of breast cancer. In a preferred embodiment, these molecules will downregulate expression and/or activity of a BSP in cells.

In another embodiment of the therapeutic methods of the present invention, a 30 pharmaceutical composition comprising a non-antibody antagonist of BSP is administered. Antagonists of BSP can be produced using methods generally known in the art. In particular, purified BSP can be used to screen libraries of pharmaceutical agents, often

combinatorial libraries of small molecules, to identify those that specifically bind and antagonize at least one activity of a BSP.

In other embodiments a pharmaceutical composition comprising an agonist of a BSP is administered. Agonists can be identified using methods analogous to those used to 5 identify antagonists.

In a preferred embodiment, the antagonist or agonist specifically binds to and antagonizes or agonizes, respectively, a BSP comprising an amino acid sequence of SEQ ID NO: 96-232, or a fusion protein, allelic variant, homolog, analog or derivative thereof. In a more preferred embodiment, the antagonist or agonist specifically binds to and 10 antagonizes or agonizes, respectively, a BSP encoded by a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1-95, or a part, allelic variant, substantially similar or hybridizing nucleic acid thereof.

Targeting Breast Tissue

15 The invention also provides a method in which a polypeptide of the invention, or an antibody thereto, is linked to a therapeutic agent such that it can be delivered to the breast or to specific cells in the breast. In a preferred embodiment, an anti-BSP antibody is linked to a therapeutic agent and is administered to a patient in need of such therapeutic agent. The therapeutic agent may be a toxin, if breast tissue needs to be selectively 20 destroyed. This would be useful for targeting and killing breast cancer cells. In another embodiment, the therapeutic agent may be a growth or differentiation factor, which would be useful for promoting breast cell function.

In another embodiment, an anti-BSP antibody may be linked to an imaging agent that can be detected using, e.g., magnetic resonance imaging, CT or PET. This would be 25 useful for determining and monitoring breast function, identifying breast cancer tumors, and identifying noncancerous breast diseases.

EXAMPLES

Example 1a: Alternative Splice Variants

We identified gene transcripts using the Gencarta™ tools (Compugen Ltd., Tel Aviv, Israel) and a variety of public and proprietary databases. These splice variants are either sequences which differ from a previously defined sequence or new uses of known 30 sequences. In general related variants are annotated as DEX0452_XXX.nt.1,

DEX0452_XXX.nt.2, DEX0452_XXX.nt.3, etc. The variant DNA sequences encode proteins which differ from a previously defined protein sequence. In relation to the nucleotide sequence naming convention, protein variants are annotated as

5 DEX0452_XXX.aa.1, DEX0452_XXX.aa.2, etc., wherein transcript DEX0452_XXX.nt.1 encodes protein DEX0452_XXX.aa.1. A single transcript may encode a protein from an alternate Open Reading Frame (ORF) which is designated DEX0452_XXX.orf.1. Additionally, multiple transcripts may encode for a single protein. In this case, DEX0452_XXX.nt.1 and DEX0452_XXX.nt.2 will both be associated with DEX0452_XXX.aa.1.

10 The mapping of the nucleic acid ("NT") SEQ ID NO; DEX ID; chromosomal location (if known); open reading frame (ORF) location; amino acid ("AA") SEQ ID NO; AA DEX ID; are shown in the table below.

SEQ ID NO	DEX ID	Chromo Map	ORF Loc	SEQ ID NO	DEX ID
1	DEX0452_001.nt.1	18q21.2	371-721	96	DEX0452_001.aa.1
2	DEX0452_002.nt.1	8q21.3	350-2305	97	DEX0452_002.aa.1
2	DEX0452_002.nt.1	8q21.3	324-2258	98	DEX0452_002.orf.1
3	DEX0452_003.nt.1	1q42.3	917-1291	99	DEX0452_003.aa.1
4	DEX0452_003.nt.2	1q42.3	1-495	100	DEX0452_003.aa.2
4	DEX0452_003.nt.2	1q42.3	1-387	101	DEX0452_003.orf.2
5	DEX0452_004.nt.1	12q14.3	2-418	102	DEX0452_004.aa.1
6	DEX0452_005.nt.1	3p21.31	151-1729	103	DEX0452_005.aa.1
6	DEX0452_005.nt.1	3p21.31	2-1156	104	DEX0452_005.orf.1
7	DEX0452_006.nt.1	1q21.1	235-1553	105	DEX0452_006.aa.1
7	DEX0452_006.nt.1	1q21.1	550-1551	106	DEX0452_006.orf.1
8	DEX0452_007.nt.1	11p15.5	357-780	107	DEX0452_007.aa.1
8	DEX0452_007.nt.1	11p15.5	447-788	108	DEX0452_007.orf.1
9	DEX0452_008.nt.1	3q26.1	263-812	109	DEX0452_008.aa.1
9	DEX0452_008.nt.1	3q26.1	252-674	110	DEX0452_008.orf.1
10	DEX0452_009.nt.1	17q12	1-396	111	DEX0452_009.aa.1
11	DEX0452_009.nt.2	17q12	644-1474	112	DEX0452_009.aa.2
12	DEX0452_010.nt.1	8q22.1	253-717	113	DEX0452_010.aa.1
13	DEX0452_011.nt.1	5q35.1	206-518	114	DEX0452_011.aa.1
13	DEX0452_011.nt.1	5q35.1	165-515	115	DEX0452_011.orf.1
14	DEX0452_012.nt.1	12q23.1	2351-3712	116	DEX0452_012.aa.1
15	DEX0452_013.nt.1	4q21.1	463-1602	117	DEX0452_013.aa.1
16	DEX0452_013.nt.2	4q21.1	34-714	118	DEX0452_013.orf.2
16	DEX0452_013.nt.2	4q21.1	33-717	119	DEX0452_013.aa.2
17	DEX0452_014.nt.1	2q35	361-663	120	DEX0452_014.aa.1
18	DEX0452_015.nt.1	15q26.2	696-1871	121	DEX0452_015.orf.1
18	DEX0452_015.nt.1	15q26.2	636-1953	122	DEX0452_015.aa.1

19	DEX0452_015.nt.2	15q26.2	361-1236	123	DEX0452_015.orf.2
19	DEX0452_015.nt.2	15q26.2	64-1317	124	DEX0452_015.aa.2
20	DEX0452_015.nt.3	15q26.2	309-1073	125	DEX0452_015.orf.3
20	DEX0452_015.nt.3	15q26.2	283-1153	126	DEX0452_015.aa.3
21	DEX0452_015.nt.4	15q26.2	106-606	127	DEX0452_015.orf.4
21	DEX0452_015.nt.4	15q26.2	120-687	128	DEX0452_015.aa.4
22	DEX0452_015.nt.5	15q26.2	3-488	129	DEX0452_015.orf.5
22	DEX0452_015.nt.5	15q26.2	118-456	130	DEX0452_015.aa.5
23	DEX0452_016.nt.1	6p21.1	1325-2242	131	DEX0452_016.orf.1
23	DEX0452_016.nt.1	6p21.1	718-2245	132	DEX0452_016.aa.1
24	DEX0452_016.nt.2	6p21.1	837-1754	133	DEX0452_016.orf.2
24	DEX0452_016.nt.2	6p21.1	466-1756	134	DEX0452_016.aa.2
25	DEX0452_016.nt.3	6p21.1	1325-2242	135	DEX0452_016.orf.3
25	DEX0452_016.nt.3	6p21.1	718-2245	132	DEX0452_016.aa.1
26	DEX0452_016.nt.4	6p21.1	1325-2242	136	DEX0452_016.orf.4
26	DEX0452_016.nt.4	6p21.1	718-2245	132	DEX0452_016.aa.1
27	DEX0452_016.nt.5	6p21.1	1325-2242	137	DEX0452_016.orf.5
27	DEX0452_016.nt.5	6p21.1	718-2245	132	DEX0452_016.aa.1
28	DEX0452_016.nt.6	6p21.1	1325-2242	138	DEX0452_016.orf.6
28	DEX0452_016.nt.6	6p21.1	718-2245	132	DEX0452_016.aa.1
29	DEX0452_017.nt.1	1q21.3	309-671	139	DEX0452_017.aa.1
30	DEX0452_018.nt.1	11q22.1	1493-1867	140	DEX0452_018.orf.1
30	DEX0452_018.nt.1	11q22.1	2980-5275	141	DEX0452_018.aa.1
31	DEX0452_019.nt.1	16p13.3	1-806	142	DEX0452_019.aa.1
31	DEX0452_019.nt.1	16p13.3	313-804	143	DEX0452_019.orf.1
32	DEX0452_020.nt.1	18q21.32	471-771	144	DEX0452_020.aa.1
32	DEX0452_020.nt.1	18q21.32	43-450	145	DEX0452_020.orf.1
33	DEX0452_021.nt.1	19q13.32	227-647	146	DEX0452_021.aa.1
33	DEX0452_021.nt.1	19q13.32	911-1405	147	DEX0452_021.orf.1
34	DEX0452_022.nt.1	7p21.1	1-408	148	DEX0452_022.aa.1
35	DEX0452_023.nt.1	8q24.13	82-669	149	DEX0452_023.aa.1
36	DEX0452_024.nt.1	3q22.1	1-212	150	DEX0452_024.aa.1
36	DEX0452_024.nt.1	3q22.1	3-209	151	DEX0452_024.orf.1
37	DEX0452_025.nt.1	2q21.2	22-548	152	DEX0452_025.aa.1
37	DEX0452_025.nt.1	2q21.2	46-546	153	DEX0452_025.orf.1
38	DEX0452_026.nt.1	14q21.1	95-469	154	DEX0452_026.aa.1
39	DEX0452_027.nt.1	5p15.33	580-897	155	DEX0452_027.aa.1
40	DEX0452_027.nt.2	5p15.33	8-718	156	DEX0452_027.aa.2
41	DEX0452_028.nt.1	5q14.3	1-206	157	DEX0452_028.aa.1
41	DEX0452_028.nt.1	5q14.3	3-470	158	DEX0452_028.orf.1
42	DEX0452_029.nt.1	12q13.12	303-2793	159	DEX0452_029.aa.1
42	DEX0452_029.nt.1	12q13.12	298-1626	160	DEX0452_029.orf.1
43	DEX0452_029.nt.2	12q13.12	450-863	161	DEX0452_029.aa.2
44	DEX0452_030.nt.1	17q12	13-196	162	DEX0452_030.aa.1
44	DEX0452_030.nt.1	17q12	487-783	163	DEX0452_030.orf.1
45	DEX0452_031.nt.1	18p11.22	169-1050	164	DEX0452_031.aa.1
46	DEX0452_031.nt.2	18p11.22	169-918	165	DEX0452_031.aa.2

47	DEX0452_031.nt.3	18p11.22	169-864	166	DEX0452_031.aa.3
48	DEX0452_032.nt.1	13	16-102	167	DEX0452_032.aa.1
48	DEX0452_032.nt.1	13	17-334	168	DEX0452_032.orf.1
49	DEX0452_033.nt.1	13	453-863	169	DEX0452_033.aa.1
50	DEX0452_033.nt.2	13	453-1175	170	DEX0452_033.aa.2
51	DEX0452_034.nt.1	17q12	1-306	171	DEX0452_034.aa.1
52	DEX0452_034.nt.2	17q12	10-635	172	DEX0452_034.aa.2
52	DEX0452_034.nt.2	17q12	8-631	173	DEX0452_034.orf.2
53	DEX0452_034.nt.3	17q12	1-309	171	DEX0452_034.aa.1
54	DEX0452_035.nt.1	16p13.3	570-1374	174	DEX0452_035.aa.1
54	DEX0452_035.nt.1	16p13.3	695-1369	175	DEX0452_035.orf.1
55	DEX0452_036.nt.1	16p13.3	579-1250	176	DEX0452_036.aa.1
56	DEX0452_036.nt.2	16p13.3	578-1481	177	DEX0452_036.aa.2
56	DEX0452_036.nt.2	16p13.3	495-1202	178	DEX0452_036.orf.2
57	DEX0452_037.nt.1	10q22.3	142-575	179	DEX0452_037.aa.1
57	DEX0452_037.nt.1	10q22.3	1-378	180	DEX0452_037.orf.1
58	DEX0452_037.nt.2	10q22.3	2-349	181	DEX0452_037.aa.2
59	DEX0452_038.nt.1	11q13.1	1-235	182	DEX0452_038.aa.1
59	DEX0452_038.nt.1	11q13.1	3063-3407	183	DEX0452_038.orf.1
60	DEX0452_038.nt.2	11q13.1	1-235	182	DEX0452_038.aa.1
60	DEX0452_038.nt.2	11q13.1	2-253	184	DEX0452_038.orf.2
61	DEX0452_038.nt.3	11q13.1	1-235	182	DEX0452_038.aa.1
61	DEX0452_038.nt.3	11q13.1	2-253	185	DEX0452_038.orf.3
62	DEX0452_039.nt.1	15q23	199-514	186	DEX0452_039.aa.1
62	DEX0452_039.nt.1	15q23	214-534	187	DEX0452_039.orf.1
63	DEX0452_040.nt.1	6p22.3	1-118	188	DEX0452_040.aa.1
63	DEX0452_040.nt.1	6p22.3	564-704	189	DEX0452_040.orf.1
64	DEX0452_041.nt.1	2q31.1	1-213	190	DEX0452_041.aa.1
65	DEX0452_042.nt.1	9q22.32	1273-1686	191	DEX0452_042.aa.1
66	DEX0452_043.nt.1	16p12.1	1-205	192	DEX0452_043.aa.1
66	DEX0452_043.nt.1	16p12.1	3-197	193	DEX0452_043.orf.1
67	DEX0452_043.nt.2	16p12.1	621-1205	194	DEX0452_043.aa.2
68	DEX0452_044.nt.1	8p11.22	29-400	195	DEX0452_044.orf.1
68	DEX0452_044.nt.1	8p11.22	88-410	196	DEX0452_044.aa.1
69	DEX0452_044.nt.2	8p11.22	3-389	197	DEX0452_044.orf.2
69	DEX0452_044.nt.2	8p11.22	2-395	198	DEX0452_044.aa.2
70	DEX0452_045.nt.1	6q22.1	915-1169	199	DEX0452_045.orf.1
70	DEX0452_045.nt.1	6q22.1	1-208	200	DEX0452_045.aa.1
71	DEX0452_046.nt.1	1q21.2	3605-4738	201	DEX0452_046.orf.1
71	DEX0452_046.nt.1	1q21.2	2985-5616	202	DEX0452_046.aa.1
72	DEX0452_046.nt.2	1q21.2	3249-4382	203	DEX0452_046.orf.2
72	DEX0452_046.nt.2	1q21.2	2913-5262	204	DEX0452_046.aa.2
73	DEX0452_047.nt.1	2p25.2	18-1364	205	DEX0452_047.aa.1
74	DEX0452_048.nt.1	18q11.2	26-1795	206	DEX0452_048.orf.1
74	DEX0452_048.nt.1	18q11.2	352-2339	207	DEX0452_048.aa.1
75	DEX0452_049.nt.1	11p15.5	905-1375	208	DEX0452_049.aa.1
76	DEX0452_049.nt.2	11p15.5	904-1378	208	DEX0452_049.aa.1

77	DEX0452_050.nt.1	11p15.2	3-809	209	DEX0452_050.aa.1
78	DEX0452_050.nt.2	11p15.2	60-1148	210	DEX0452_050.aa.2
79	DEX0452_051.nt.1	11p15.2	251-1510	211	DEX0452_051.aa.1
80	DEX0452_052.nt.1	5q13.3	323-808	212	DEX0452_052.aa.1
81	DEX0452_053.nt.1	10q26.12	527-733	213	DEX0452_053.orf.1
81	DEX0452_053.nt.1	10q26.12	1-130	214	DEX0452_053.aa.1
82	DEX0452_054.nt.1	X;115879825 -115903932	1-516	215	DEX0452_054.orf.1
82	DEX0452_054.nt.1	X;115879825 -115903932	115-520	216	DEX0452_054.aa.1
83	DEX0452_055.nt.1	1q23.1	217-1404	217	DEX0452_055.aa.1
84	DEX0452_055.nt.2	1q23.1	857-1621	218	DEX0452_055.aa.2
85	DEX0452_056.nt.1	8q22.3	1358-2593	219	DEX0452_056.orf.1
85	DEX0452_056.nt.1	8q22.3	1-171	220	DEX0452_056.aa.1
86	DEX0452_057.nt.1	10q26.13	337-1626	221	DEX0452_057.orf.1
86	DEX0452_057.nt.1	10q26.13	471-1629	222	DEX0452_057.aa.1
87	DEX0452_058.nt.1	4q25	92-460	223	DEX0452_058.aa.1
88	DEX0452_058.nt.2	1q23.1	1443-2075	224	DEX0452_058.orf.2
88	DEX0452_058.nt.2	1q23.1	1541-2075	225	DEX0452_058.aa.2
89	DEX0452_058.nt.3	1q23.1	1023-1557	225	DEX0452_058.aa.2
89	DEX0452_058.nt.3	1q23.1	925-1557	226	DEX0452_058.orf.3
90	DEX0452_058.nt.4	1q23.1	895-1430	225	DEX0452_058.aa.2
90	DEX0452_058.nt.4	1q23.1	798-1430	227	DEX0452_058.orf.4
91	DEX0452_058.nt.5	1q23.1	731-1265	225	DEX0452_058.aa.2
91	DEX0452_058.nt.5	1q23.1	633-1265	228	DEX0452_058.orf.5
92	DEX0452_058.nt.6	1q23.1	872-1406	225	DEX0452_058.aa.2
92	DEX0452_058.nt.6	1q23.1	774-1406	229	DEX0452_058.orf.6
93	DEX0452_058.nt.7	1q23.1	907-1441	225	DEX0452_058.aa.2
93	DEX0452_058.nt.7	1q23.1	809-1441	230	DEX0452_058.orf.7
94	DEX0452_058.nt.8	1q23.1	528-1062	225	DEX0452_058.aa.2
94	DEX0452_058.nt.8	1q23.1	430-1062	231	DEX0452_058.orf.8
95	DEX0452_058.nt.9	1q23.1	402-937	225	DEX0452_058.aa.2
95	DEX0452_058.nt.9	1q23.1	305-937	232	DEX0452_058.orf.9

The polypeptides of the present invention were analyzed and the following attributes were identified; specifically, epitopes, post translational modifications, signal peptides and transmembrane domains. Antigenicity (Epitope) prediction was performed through the antigenic module in the EMBOSS package. Rice, P., EMBOSS: The European Molecular Biology Open Software Suite, *Trends in Genetics* 16(6): 276-277 (2000). The antigenic module predicts potentially antigenic regions of a protein sequence, using the method of Kolaskar and Tongaonkar. Kolaskar, AS and Tongaonkar, PC., A semi-empirical method for prediction of antigenic determinants on protein antigens, *FEBS Letters* 276: 172-174 (1990). Examples of post-translational modifications (PTMs) and other motifs of the BSPs of this invention are listed below. In addition, antibodies that

- specifically bind such post-translational modifications may be useful as a diagnostic or as therapeutic. The PTMs and other motifs were predicted by using the ProSite Dictionary of Proteins Sites and Patterns (Bairoch *et al.*, *Nucleic Acids Res.* 25(1):217-221 (1997)), the following motifs, including PTMs, were predicted for the BSPs of the invention. The
- 5 signal peptides were detected by using the SignalP 2.0, *see* Nielsen *et al.*, *Protein Engineering* 12, 3-9 (1999). Prediction of transmembrane helices in proteins was performed by the application TMHMM 2.0, "currently the best performing transmembrane prediction program", according to authors (Krogh *et al.*, *Journal of Molecular Biology*, 305(3):567-580, (2001); Moller *et al.*, *Bioinformatics*, 17(7):646-653, (2001);
- 10 Sonnhammer, *et al.*, *A hidden Markov model for predicting transmembrane helices in protein sequences* in Glasgow, *et al.* Ed. Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology, pages 175-182, Menlo Park, CA, 1998. AAAI Press. The PSORT II program may also be used to predict cellular localizations. Horton *et al.*, *Intelligent Systems for Molecular Biology* 5: 147-152 (1997).
- 15 The table below includes the following sequence annotations: Signal peptide presence; TM (number of membrane domain, topology in orientation and position); Amino acid location and antigenic index (location, AI score); PTM and other motifs (type, amino acid residue locations); and functional domains (type, amino acid residue locations).

DEX ID	Sig P	TMHMM	Antigenicity	PTM	Domains
DEX0452_001.aa.1	N	0 - o1- 117;	29-35,1.089; 89-97,1.091; 7-16,1.157; 37-45,1.077; 53-67,1.106; 104- 114,1.158;	MYRISTYL 91-96;	RASSTRNSFRMNG 112-117; RASSTRNSFRMNG 59-72; RAB 8- 112; RAS 5-112;
DEX0452_002.aa.1	N	0 - o1- 651;	67-73,1.096; 553- 561,1.091; 364- 370,1.067; 90-109,1.13; 439- 448,1.082; 151- 158,1.082; 191- 209,1.254; 216- 223,1.088; 134- 146,1.198; 455-	MYRISTYL 575-580; PKC_PHOSPHO_SITE 15-17; CK2_PHOSPHO_SITE 76-79; ASN_GLYCOSYLATION 648- 651; MYRISTYL 582-587; PKC_PHOSPHO_SITE 87-89; CK2_PHOSPHO_SITE 585- 588; PKC_PHOSPHO_SITE 26-28; CAMP_PHOSPHO_SITE 535- 538; CK2_PHOSPHO_SITE 597-600; PKC_PHOSPHO_SITE 416- 418; PKC_PHOSPHO_SITE 644-646; CK2_PHOSPHO_SITE 322- 325; PKC PHOSPHO SITE	WWDOMAIN 351- 364; WW_DOMAIN_1 387-412; C2_DOMAIN_2 5- 98; WW 350-382; WW_DOMAIN_1 462-487; WW 458-487; C2 19- 113; WW 383- 414; WW 457- 489; C2 20-98; WWDOMAIN 473- 487; WW_DOMAIN_1 355-380; WW 383-412;

			460, 1.077; 77-85, 1.071; 295- 300, 1.089; 380- 385, 1.056; 41-48, 1.175; 230- 236, 1.071; 336- 341, 1.063; 308- 313, 1.038; 525- 532, 1.072; 113- 122, 1.087; 469- 476, 1.11; 279- 288, 1.111; 258- 266, 1.129; 396- 402, 1.133; 504- 523, 1.201; 17-28, 1.172; 623- 634, 1.104;	53-55; ASN_GLYCOSYLATION 647- 650; CK2_PHOSPHO_SITE 309-312; PKC_PHOSPHO_SITE 5-7; PKC_PHOSPHO_SITE 550- 552; CK2_PHOSPHO_SITE 353-356; CK2_PHOSPHO_SITE 276- 279; ASN_GLYCOSYLATION 640-643; PKC_PHOSPHO_SITE 618- 620; PKC_PHOSPHO_SITE 469-471; CK2_PHOSPHO_SITE 288- 291; ASN_GLYCOSYLATION 254-257; CK2_PHOSPHO_SITE 482- 485; ASN_GLYCOSYLATION 151-154; PKC_PHOSPHO_SITE 425- 427; CK2_PHOSPHO_SITE 375-378; ASN_GLYCOSYLATION 148- 151; PKC_PHOSPHO_SITE 178-180; ASN_GLYCOSYLATION 13- 16; CK2_PHOSPHO_SITE 277-280; CK2_PHOSPHO_SITE 299- 302; PKC_PHOSPHO_SITE 174-176; MYRISTYL 128- 133; MYRISTYL 344-349; PKC_PHOSPHO_SITE 509- 511; ASN_GLYCOSYLATION 272-275; AMIDATION 498- 501; PKC_PHOSPHO_SITE 84-86; MYRISTYL 543- 548; ASN_GLYCOSYLATION 607-610; CK2_PHOSPHO_SITE 618- 621;	WW_DOMAIN_2_1 349-382; WW 351-380; WW_DOMAIN_2_3 456-489; WW_DOMAIN_2_2 381-414;
DEX0452_002.orf. 1	N	0 - 645;	50-57, 1.175; 288- 297, 1.111; 448- 457, 1.082; 76-82, 1.096; 200- 218, 1.254; 478- 485, 1.11; 122- 131, 1.087; 405- 411, 1.133; 239- 245, 1.071; 513-	PKC_PHOSPHO_SITE 478- 480; ASN_GLYCOSYLATION 22-25; MYRISTYL 591- 596; CK2_PHOSPHO_SITE 318-321; CK2_PHOSPHO_SITE 308- 311; PKC_PHOSPHO_SITE 425-427; ASN_GLYCOSYLATION 157- 160; PKC_PHOSPHO_SITE 96-98; MYRISTYL 552- 557; ASN_GLYCOSYLATION 616-619; MYRISTYL 353- 358; PKC_PHOSPHO_SITE 187-189; CK2_PHOSPHO_SITE 491- 494; PKC_PHOSPHO SITE	C2 28-122; WW_DOMAIN_1 396-421; WW 466-498; WW 467-496; WW 392-421; C2 29- 107; WW 359- 391; WW_DOMAIN_1 471-496; WW 360-389; WW 392-423; WWDOMAIN 482- 496; WWDOMAIN 360-373; WW_DOMAIN_2_3 465-498;

			532, 1.201; 86-94, 1.071; 373- 379, 1.067; 534- 541, 1.072; 26-37, 1.172; 160- 167, 1.082; 464- 469, 1.077; 304- 309, 1.089; 99-118, 1.13; 562- 570, 1.091; 317- 322, 1.038; 389- 394, 1.056; 225- 232, 1.088; 345- 350, 1.063; 143- 155, 1.198; 267- 275, 1.129;	434-436; CK2_PHOSPHO_SITE 594- 597; PKC_PHOSPHO_SITE 93-95; CK2_PHOSPHO_SITE 606-609; CK2_PHOSPHO_SITE 286- 289; CK2_PHOSPHO_SITE 384-387; PKC_PHOSPHO_SITE 62-64; PKC_PHOSPHO_SITE 518- 520; PKC_PHOSPHO_SITE 35-37; AMIDATION 507- 510; ASN_GLYCOSYLATION 281-284; CK2_PHOSPHO_SITE 331- 334; MYRISTYL 137-142; PKC_PHOSPHO_SITE 14-16; CAMP_PHOSPHO_SITE 544- 547; PKC_PHOSPHO_SITE 559-561; CK2_PHOSPHO_SITE 297- 300; CK2_PHOSPHO_SITE 362-365; PKC_PHOSPHO_SITE 183- 185; ASN_GLYCOSYLATION 263-266; CK2_PHOSPHO_SITE 285- 288; CK2_PHOSPHO_SITE 627-630; CK2_PHOSPHO_SITE 85-88; PKC_PHOSPHO_SITE 24-26; PKC_PHOSPHO_SITE 627- 629; MYRISTYL 584-589; ASN_GLYCOSYLATION 160- 163;	C2_DOMAIN_2 14- 107; WW_DOMAIN_2_1 358-391; WW_DOMAIN_2_2 390-423; WW_DOMAIN_1 364-389;
DEX0452_003.aa.1	N	0 - o1- 125;	87-98, 1.08; 5-20, 1.109; 55-60, 1.073; 27-48, 1.145; 76-85, 1.088;	PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 84-87; PKC_PHOSPHO_SITE 51-53; CK2_PHOSPHO_SITE 51-54; PKC_PHOSPHO_SITE 122- 124;	
DEX0452_003.aa.2	N	0 - o1- 164;	29-37, 1.176; 42- 144, 1.183; 4-17, 1.167;	MYRISTYL 97-102; AMIDATION 90-93; MYRISTYL 22-27; MYRISTYL 24-29;	
DEX0452_003.orf.2	N	0 - o1- 129;	42-117, 1.17; 119- 126, 1.202; 4-17, 1.167; 29-37, 1.176;	MYRISTYL 22-27; MYRISTYL 24-29; AMIDATION 90-93; MYRISTYL 97-102;	
DEX0452_004.aa.1	N	3 - o1- 51;tm 52- 74;i7 5- 80;tm 81-	78- 136, 1.168; 4-20, 1.197; 39-76, 1.208; 26-36, 1.167;	PKC_PHOSPHO_SITE 21-23; LEUCINE_ZIPPER 106-127; CK2_PHOSPHO_SITE 107- 110;	

		98;09 9- 110;t m111- 133;i 134- 139;			
DEX0452_005.aa.1	N	0 - 01- 525;	286- 337,1.143; 436- 448,1.138; 4-23,1.127; 103- 129,1.159; 465- 485,1.151; 382- 405,1.131; 212- 226,1.235; 132- 157,1.17; 276- 284,1.128; 512- 520,1.157; 47-55,1.098; 496- 508,1.194; 411- 419,1.138; 346- 380,1.113; 83-95,1.193; 58-79,1.138; 487- 493,1.075; 252- 274,1.143; 181- 209,1.169; 244- 250,1.061;	CK2_PHOSPHO_SITE 133- 136; CAMP_PHOSPHO_SITE 488-491; CK2_PHOSPHO_SITE 447- 450; CK2_PHOSPHO_SITE 449-452; CK2_PHOSPHO_SITE 70-73; CAMP_PHOSPHO_SITE 143- 146; MYRISTYL 162-167; CK2_PHOSPHO_SITE 229- 232; CK2_PHOSPHO_SITE 384-387; CK2_PHOSPHO_SITE 53-56; CK2_PHOSPHO_SITE 256- 259; PKC_PHOSPHO_SITE 229-231; CK2_PHOSPHO_SITE 129- 132; PKC_PHOSPHO_SITE 463-465; CK2_PHOSPHO_SITE 403- 406; PKC_PHOSPHO_SITE 487-489;	PRICHEXTENSN 345-361; PRICHEXTENSN 295-311; zf- MYND 479-515; PRO_RICH 295- 361;
DEX0452_005.orf.1	N	0 - 01- 385;	231- 259,1.169; 326- 334,1.128; 262- 276,1.235; 294- 300,1.061; 336- 382,1.143; 108- 129,1.138; 182- 207,1.17; 11-44,1.14; 133-	CK2_PHOSPHO_SITE 183- 186; CK2_PHOSPHO_SITE 279-282; MYRISTYL 48- 53; PKC_PHOSPHO_SITE 279-281; CK2_PHOSPHO_SITE 179- 182; CAMP_PHOSPHO_SITE 193-196; CK2_PHOSPHO_SITE 306- 309; CK2_PHOSPHO_SITE 120-123; MYRISTYL 212- 217; CK2_PHOSPHO_SITE 103-106;	

			145, 1.193; 54-73, 1.127; 153- 179, 1.159; 302- 324, 1.143; 97- 105, 1.098;		
DEX0452_006.aa.1	N	0 - o1- 438;	275- 321, 1.134; 420- 435, 1.158; 184- 193, 1.106; 259- 268, 1.098; 90- 152, 1.166; 15-24, 1.098; 31-77, 1.134; 353- 366, 1.124; 386- 418, 1.117; 200- 245, 1.12; 368- 384, 1.092;	AMIDATION 159-162; CK2_PHOSPHO_SITE 122- 125; CK2_PHOSPHO_SITE 64-67; AMIDATION 164- 167; CK2_PHOSPHO_SITE 308-311; MYRISTYL 344- 349; CAMP_PHOSPHO_SITE 342-345; PKC_PHOSPHO_SITE 168- 170; CK2_PHOSPHO_SITE 402-405; MYRISTYL 55- 60; CK2_PHOSPHO_SITE 337-340; PKC_PHOSPHO_SITE 337- 339; AMIDATION 174-177; AMIDATION 333-336; MYRISTYL 299-304; AMIDATION 328-331; CK2_PHOSPHO_SITE 168- 171;	
DEX0452_006.orf.	Y	0 - o1- 1 334;	171- 217, 1.134; 264- 280, 1.092; 249- 262, 1.124; 96-141, 1.12; 80-89, 1.106; 316- 331, 1.158; 282- 314, 1.117; 155- 164, 1.098; 4-26, 1.228; 41-48, 1.101; 28-34, 1.042;	CK2_PHOSPHO_SITE 233- 236; MYRISTYL 240-245; CK2_PHOSPHO_SITE 27-30; AMIDATION 224-227; AMIDATION 70-73; AMIDATION 55-58; PKC_PHOSPHO_SITE 233- 235; CK2_PHOSPHO_SITE 204-207; CAMP_PHOSPHO_SITE 238- 314; MYRISTYL 195-200; PKC_PHOSPHO_SITE 1-3; CK2_PHOSPHO_SITE 298- 301; PKC_PHOSPHO_SITE 4-6; AMIDATION 229-232;	
DEX0452_007.aa.1	N	1 - i1- 51;tm 52- 74;o7 5- 140;	17-29, 1.128; 41-81, 1.21; 91- 137, 1.189;	MYRISTYL 60-65; ASN GLYCOSYLATION 85- 88; PKC_PHOSPHO_SITE 27-29; PKC_PHOSPHO_SITE 7-9; PKC_PHOSPHO_SITE 48-50; CAMP_PHOSPHO_SITE 2-5; PKC_PHOSPHO_SITE 87-89; MYRISTYL 91-96;	CD225 4-68;
DEX0452_007.orf.	N	0 - i1- 114;		TYR_PHOSPHO_SITE 43-50; MYRISTYL 1-6; MYRISTYL 60-65; PKC_PHOSPHO_SITE 56-58; PKC_PHOSPHO SITE	CYTOCHROME_C 103-108;

				108-110; MYRISTYL 41-46; ASN_GLYCOSYLATION 94-97;	
DEX0452_008.aa.1	N	0 - o1- 182;	64-71, 1.073; 105- 114, 1.096; 135- 143, 1.074; 157- 179, 1.123; 31-46, 1.081; 14-29, 1.149; 120- 132, 1.096;	CK2_PHOSPHO_SITE 143-146; PKC_PHOSPHO_SITE 147-149; PKC_PHOSPHO_SITE 126-128; ASN_GLYCOSYLATION 33-36; ASN_GLYCOSYLATION 71-74; PKC_PHOSPHO_SITE 177-179; PKC_PHOSPHO_SITE 39-41;	
DEX0452_008.orf.1	N	0 - il- 141;	124- 134, 1.096; 109- 118, 1.096; 18-33, 1.149; 35-50, 1.081; 68-75, 1.073;	PKC_PHOSPHO_SITE 136-138; ASN_GLYCOSYLATION 37-40; CK2_PHOSPHO_SITE 4-7; PKC_PHOSPHO_SITE 130-132; PKC_PHOSPHO_SITE 43-45; ASN_GLYCOSYLATION 75-78;	
DEX0452_009.aa.1	N	0 - o1- 132;	4-19, 1.181; 43-66, 1.143; 113- 120, 1.118; 122- 129, 1.155; 71- 111, 1.241; 22-34, 1.14;	CK2_PHOSPHO_SITE 61-64; PKC_PHOSPHO_SITE 37-39; CK2_PHOSPHO_SITE 15-18; CK2_PHOSPHO_SITE 70-73;	
DEX0452_009.aa.2	N	0 - o1- 277;	90-97, 1.055; 143- 156, 1.179; 124- 139, 1.173; 100- 110, 1.201; 41-49, 1.115; 161- 173, 1.163; 187- 211, 1.143; 258- 265, 1.118; 4-22, 1.108; 113- 121, 1.14; 216- 256, 1.241; 178- 183, 1.036; 62-87, 1.118; 267- 274, 1.155;	PKC_PHOSPHO_SITE 27-29; PKC_PHOSPHO_SITE 123-125; CK2_PHOSPHO_SITE 13-16; CAMP_PHOSPHO_SITE 59-62; CK2_PHOSPHO_SITE 215-218; AMIDATION 50-53; MYRISTYL 111-116; MYRISTYL 85-90; MYRISTYL 22-27; PKC_PHOSPHO_SITE 50-52; CK2_PHOSPHO_SITE 12-15; LEUCINE_ZIPPER 131-152; MYRISTYL 144-149; CK2_PHOSPHO_SITE 163-166; PKC_PHOSPHO_SITE 89-91; CK2_PHOSPHO_SITE 206-209;	PRICHEXTENSN 61-82; RA 100-186; PRICHEXTENSN 239-251; PRICHEXTENSN 6-18; RA 100-186; PRICHEXTENSN 22-38; PRO_RICH 7-87; RA_DOMAIN 100-186;
DEX0452_010.aa.1	Y	0 - o1- 155;	121- 133, 1.062; 4-70, 1.215;	MYRISTYL 84-89; ASN_GLYCOSYLATION 97-100; MYRISTYL 113-118;	

			76-83, 1.084; 88-96, 1.059; 98- 113, 1.099;	ASN_GLYCOSYLATION 74-77; MYRISTYL 137-142;	
DEX0452_011.aa.1	N	1 - o1- 69;tm 70- 92;i9 3- 103;	23-49, 1.153; 15-20, 1.034; 62- 100, 1.293;	PKC_PHOSPHO_SITE 9-11; CK2_PHOSPHO_SITE 48-51; CK2_PHOSPHO_SITE 29-32; CK2_PHOSPHO_SITE 17-20;	
DEX0452_011.orf.1	N	1 - o1- 83;tm 84- 106;i 107- 117;		AMIDATION 8-11; PKC_PHOSPHO_SITE 23-25; CK2_PHOSPHO_SITE 43-46; CK2_PHOSPHO_SITE 31-34; CK2_PHOSPHO_SITE 62-65; PKC_PHOSPHO_SITE 8-10;	
DEX0452_012.aa.1	N			PKC_PHOSPHO_SITE 211-213; CK2_PHOSPHO_SITE 229-232; PKC_PHOSPHO_SITE 180-182; PKC_PHOSPHO_SITE 237-239; 297, 1.151; 439- 450, 1.101; 359- 366, 1.064; 207- 216, 1.073; 143- 150, 1.089; 17-29, 1.106; 7-14, 1.097; 1 - 74-82, 1.079; o1- 372- 410;t m411- 270- 430;i 279, 1.072; 431- 159- 454; 167, 1.066; 121- 138, 1.142; 34-45, 1.229; 304- 312, 1.108; 88-93, 1.036; 410- 433, 1.277; 240- 245, 1.061; 387- 401, 1.176;	PKC_PHOSPHO_SITE 93-96; CK2_PHOSPHO_SITE 190-193; PKC_PHOSPHO_SITE 362-364; ASN_GLYCOSYLATION 178-181; MYRISTYL 80-85; CK2_PHOSPHO_SITE 116-119; CK2_PHOSPHO_SITE 66-69; CK2_PHOSPHO_SITE 167-170; ASN_GLYCOSYLATION 24-27; MYRISTYL 78-83; ASN_GLYCOSYLATION 175-178; AMIDATION 91-94; CK2_PHOSPHO_SITE 184-187; ASN_GLYCOSYLATION 220-223; PKC_PHOSPHO_SITE 402-404; PKC_PHOSPHO_SITE 405-407; MYRISTYL 236-241; PKC_PHOSPHO_SITE 46-48; ASN_GLYCOSYLATION 290-293; PKC_PHOSPHO_SITE 257-259; CK2_PHOSPHO_SITE 67-70; PKC_PHOSPHO_SITE 137-139; MYRISTYL 176-181; PKC_PHOSPHO SITE 99- LEM 109-152; LEM 110-153;

				101; PKC_PHOSPHO_SITE 387-389; MYRISTYL 227- 232; MYRISTYL 358-363; PKC_PHOSPHO_SITE 96-98; PKC_PHOSPHO_SITE 254- 256;	
DEX0452_013.aa.1	N	0 - o1- 380;	215- 269,1.266; 5-38,1.277; 280- 311,1.186; 182- 189,1.076; 317- 329,1.126; 96- 152,1.214; 42-62,1.182; 86-91,1.09; 354- 359,1.061; 158- 172,1.143; 73-79,1.07; 369- 377,1.161; 194- 210,1.163; 361- 367,1.066;	PKC_PHOSPHO_SITE 309- 311; ASN,GLYCOSYLATION 273-276; PKC_PHOSPHO_SITE 182- 184; CK2_PHOSPHO_SITE 338-341; CK2_PHOSPHO_SITE 292- 295; TYR_PHOSPHO_SITE 179-187; LEUCINE_ZIPPER 8-29; CK2_PHOSPHO_SITE 322-325; MYRISTYL 90- 95; PKC_PHOSPHO_SITE 210-212; CK2_PHOSPHO_SITE 343- 346; PKC_PHOSPHO_SITE 253-255; CAMP_PHOSPHO_SITE 267- 270; CK2_PHOSPHO_SITE 328-331; CK2_PHOSPHO_SITE 170- 173; CK2_PHOSPHO_SITE 75-78; CK2_PHOSPHO_SITE 348-351; CK2_PHOSPHO_SITE 108- 111; CK2_PHOSPHO_SITE 360-363; CK2_PHOSPHO_SITE 157- 160;	cyclin 55-190; CYCLIN 97-183;
DEX0452_013.orf.	N	0 - o1- 2	227;	TYR_PHOSPHO_SITE 26-34; CAMP_PHOSPHO_SITE 114- 117; CK2_PHOSPHO_SITE 17-20; CK2_PHOSPHO_SITE 185-188; CK2_PHOSPHO_SITE 6-9; CK2_PHOSPHO_SITE 175- 178; PKC_PHOSPHO_SITE 11-13; MYRISTYL 6-11; PKC_PHOSPHO_SITE 156- 158; CK2_PHOSPHO_SITE 207-210; PKC_PHOSPHO_SITE 57-59; MYRISTYL 10-15; PKC_PHOSPHO_SITE 29-31; CK2_PHOSPHO_SITE 195- 198; CK2_PHOSPHO_SITE 169-172; PKC_PHOSPHO_SITE 100- 102; ASN,GLYCOSYLATION 120-123; AMIDATION 10- 13; AMIDATION 9-12; CK2_PHOSPHO_SITE 190- 193; CK2_PHOSPHO_SITE 139-142;	

DEX0452_013.aa.2	N	0 - o1- 227;	62- 116, 1.266; 41-57, 1.163; 127- 158, 1.186; 201- 206, 1.061; 164- 176, 1.126; 29-36, 1.076; 216- 224, 1.161; 208- 214, 1.066;	CK2_PHOSPHO_SITE 169-172; ASN,GLYCOSYLATION 120-123; TYR_PHOSPHO_SITE 26-34; CK2_PHOSPHO_SITE 139-142; CAMP_PHOSPHO_SITE 114-117; CK2_PHOSPHO_SITE 175-178; CK2_PHOSPHO_SITE 17-20; MYRISTYL 10-15; PKC_PHOSPHO_SITE 57-59; CK2_PHOSPHO_SITE 195-198; PKC_PHOSPHO_SITE 100-102; CK2_PHOSPHO_SITE 190-193; CK2_PHOSPHO_SITE 207-210; CK2_PHOSPHO_SITE 185-188; PKC_PHOSPHO_SITE 156-158; PKC_PHOSPHO_SITE 29-31;	
DEX0452_014.aa.1	Y	1 - il- 61;tm 62- 84;o8 5- 101;	36-46, 1.134; 8-25, 1.122; 48-95, 1.27;	CK2_PHOSPHO_SITE 21-24; MICROBODIES_CTER 99-101; PKC_PHOSPHO_SITE 93-95; MYRISTYL 67-72; MYRISTYL 87-92;	
DEX0452_015.orf.1	Y	0 - o1- 392;	337- 346, 1.179; 169- 188, 1.13; 4- 25, 1.199; 355- 365, 1.104; 231- 250, 1.148; 35-56, 1.192; 256- 264, 1.106; 288- 294, 1.044; 121- 128, 1.058; 269- 279, 1.202; 132- 138, 1.042; 88- 104, 1.094;	CK2_PHOSPHO_SITE 383-386; PKC_PHOSPHO_SITE 137-139; MYRISTYL 357-362; PKC_PHOSPHO_SITE 131-133; CK2_PHOSPHO_SITE 137-140; MYRISTYL 156-161; MYRISTYL 160-165; PKC_PHOSPHO_SITE 78-80; PKC_PHOSPHO_SITE 380-382; MYRISTYL 268-273; ASN,GLYCOSYLATION 76-79; PKC_PHOSPHO_SITE 189-191; MYRISTYL 77-82; ASN,GLYCOSYLATION 373-376; CK2_PHOSPHO_SITE 216-219; PKC_PHOSPHO_SITE 44-46; CK2_PHOSPHO_SITE 57-60;	isodh 9-381; IDH_IMDH 250-269; nadp_idh_euk 9-391;
DEX0452_015.aa.1	Y	0 - o1- 438;	108- 124, 1.094; 375- 385, 1.104; 189- 208, 1.13; 55-76, 1.192; 308- 314, 1.044;	MYRISTYL 176-181; PKC_PHOSPHO_SITE 157-159; ASN,GLYCOSYLATION 96-99; CK2_PHOSPHO_SITE 77-80; MYRISTYL 410-415; MYRISTYL 420-425; CK2_PHOSPHO_SITE 403-406; MYRISTYL 414-419; PKC_PHOSPHO SITE 400-	nadp_idh_euk 29-411; isodh 19-401; IDH_IMDH 270-289;

			4-15, 1.225; 17-45, 1.22; 152- 158, 1.042; 289- 299, 1.202; 251- 270, 1.148; 276- 284, 1.106; 357- 366, 1.179; 141- 148, 1.058;	402; MYRISTYL 416-421; MYRISTYL 97-102; PKC_PHOSPHO_SITE 209- 211; CK2_PHOSPHO_SITE 157-160; MYRISTYL 423- 428; MYRISTYL 377-382; MYRISTYL 180-185; PKC_PHOSPHO_SITE 151- 153; ASN_GLYCOSYLATION 393-396; MYRISTYL 288- 293; PKC_PHOSPHO_SITE 64-66; CK2_PHOSPHO_SITE 236-239; PKC_PHOSPHO_SITE 98- 100;	
DEX0452_015.orf.	N	0 - 01- 292;	237- 246, 1.179; 131- 150, 1.148; 255- 265, 1.104; 32-38, 1.042; 69-88, 1.13; 8-28, 1.175; 156- 164, 1.106; 169- 179, 1.202; 188- 194, 1.044;	ASN_GLYCOSYLATION 273- 276; MYRISTYL 56-61; CK2_PHOSPHO_SITE 37-40; CK2_PHOSPHO_SITE 116- 119; CK2_PHOSPHO_SITE 283-286; PKC_PHOSPHO_SITE 280- 282; PKC_PHOSPHO_SITE 37-39; MYRISTYL 60-65; MYRISTYL 168-173; MYRISTYL 257-262; PKC_PHOSPHO_SITE 31-33; PKC_PHOSPHO_SITE 89-91;	nadp_idh_euk 18-291; isodh 1-281; IDH_IMDH 150-169;
DEX0452_015_aa.2	Y	0 - 01- 417;	168- 187, 1.13; 131- 137, 1.042; 27-37, 1.11; 255- 263, 1.106; 268- 278, 1.202; 230- 249, 1.148; 287- 293, 1.044; 7-19, 1.287; 354- 364, 1.104; 107- 127, 1.175; 336- 345, 1.179; 59-75, 1.094;	ASN_GLYCOSYLATION 372- 375; MYRISTYL 393-398; MYRISTYL 23-28; ASN_GLYCOSYLATION 47- 50; PKC_PHOSPHO_SITE 49-51; CK2_PHOSPHO_SITE 382-385; MYRISTYL 356- 361; PKC_PHOSPHO_SITE 130-132; CK2_PHOSPHO_SITE 215- 218; MYRISTYL 389-394; MYRISTYL 395-400; MYRISTYL 24-29; MYRISTYL 159-164; PKC_PHOSPHO_SITE 136- 138; PKC_PHOSPHO_SITE 188-190; MYRISTYL 155- 160; MYRISTYL 402-407; PKC_PHOSPHO_SITE 379- 381; MYRISTYL 48-53; CK2_PHOSPHO_SITE 136- 139; MYRISTYL 267-272; MYRISTYL 399-404;	isodh 1-380; nadp_idh_euk 117-390; IDH_IMDH 249- 268;
DEX0452_015.orf.	Y	0 - 01- 255;	119- 127, 1.106; 94- 113, 1.148; 57-71, 1.163; 200-	PKC_PHOSPHO_SITE 73-75; MYRISTYL 220-225; CK2_PHOSPHO_SITE 42-45; ASN_GLYCOSYLATION 59- 62; MYRISTYL 55-60; PKC PHOSPHO SITE 243-	isodh 1-244; IDH_IMDH 113- 132;

			209, 1.179; 218- 228, 1.104; 20-52, 1.194; 151- 157, 1.044; 132- 142, 1.202;	245; CK2_PHOSPHO_SITE 79-82; PKC_PHOSPHO_SITE 10-12; CK2_PHOSPHO_SITE 246-249; PKC_PHOSPHO_SITE 11-13; MYRISTYL 131-136; ASN_GLYCOSYLATION 236- 239; PKC_PHOSPHO_SITE 16-18; MYRISTYL 5-10;	
DEX0452_015.aa.3	N	0 - o1- 289;	159- 165, 1.044; 127- 135, 1.106; 226- 236, 1.104; 102- 121, 1.148; 29-60, 1.194; 140- 150, 1.202; 65-79, 1.163; 208- 217, 1.179;	MYRISTYL 274-279; MYRISTYL 139-144; CK2_PHOSPHO_SITE 87-90; MYRISTYL 267-272; MYRISTYL 228-233; CK2_PHOSPHO_SITE 50-53; MYRISTYL 271-276; ASN_GLYCOSYLATION 67- 70; MYRISTYL 261-266; PKC_PHOSPHO_SITE 251- 253; PKC_PHOSPHO_SITE 81-83; PKC_PHOSPHO_SITE 2-4; ASN_GLYCOSYLATION 244-247; CK2_PHOSPHO_SITE 254- 257; MYRISTYL 63-68; MYRISTYL 265-270;	IDH_IMDH 121- 140; isodh 1- 252;
DEX0452_015.orf.4	N	0 - o1- 167;	63-69, 1.044; 8-14, 1.104; 44-54, 1.202; 32-42, 1.058; 112- 121, 1.179; 130- 140, 1.104;	ASN_GLYCOSYLATION 148- 151; MYRISTYL 15-20; PKC_PHOSPHO_SITE 155- 157; MYRISTYL 132-137; MYRISTYL 43-48; CK2_PHOSPHO_SITE 158- 161;	
DEX0452_015.aa.4	N	0 - o1- 188;	27-37, 1.058; 125- 135, 1.104; 4-9, 1.104; 58-64, 1.044; 39-49, 1.202; 107- 116, 1.179;	ASN_GLYCOSYLATION 143- 146; MYRISTYL 173-178; PKC_PHOSPHO_SITE 150- 152; MYRISTYL 160-165; MYRISTYL 170-175; MYRISTYL 127-132; CK2_PHOSPHO_SITE 153- 156; MYRISTYL 10-15; MYRISTYL 38-43; MYRISTYL 166-171; MYRISTYL 164-169;	
DEX0452_015.orf.5	N	0 - o1- 162;	4-16, 1.133; 43-54, 1.169; 82-92, 1.162; 59-68, 1.068; 26-41, 1.156; 111- 127, 1.071;	MYRISTYL 130-135; MYRISTYL 138-143; PKC_PHOSPHO_SITE 14-16; MYRISTYL 144-149; CK2_PHOSPHO_SITE 69-72; MYRISTYL 147-152; MYRISTYL 55-60; CK2_PHOSPHO_SITE 19-22; MYRISTYL 140-145; MYRISTYL 134-139;	
DEX0452_015.aa.5	N	0 - o1- 112;	4-15, 1.169; 72-88, 1.071; 43-53, 1.162;	MYRISTYL 16-21; AMIDATION 99-102; CK2_PHOSPHO_SITE 30-33;	

			20-29, 1.068;		
DEX0452_016.orf.	N	0 - o1- 1 306;	PKC_PHOSPHO_SITE 291-293; PKC_PHOSPHO_SITE 80-82; MYRISTYL 249-254; CK2_PHOSPHO_SITE 48-51; 282- 299, 1.158; 163- 177, 1.099; 109- 136, 1.193; 274- 280, 1.052; 229- 267, 1.232; 75-84, 1.204; 9-42, 1.246; 193- 223, 1.135; 87-107, 1.22; 141- 159, 1.222; 62-67, 1.054;	CK2_PHOSPHO_SITE 102-105; CK2_PHOSPHO_SITE 108-111; PKC_PHOSPHO_SITE 273-275; MYRISTYL 283-288; CK2_PHOSPHO_SITE 173-176; CK2_PHOSPHO_SITE 298-301; CK2_PHOSPHO_SITE 80-83; CK2_PHOSPHO_SITE 278-281; PKC_PHOSPHO_SITE 266-268; PKC_PHOSPHO_SITE 278-280;	CYTOCHROME_C 248-253;
DEX0452_016.aa.1	N	0 - o1- 1 508;	395- 425, 1.135; 484- 501, 1.158; 212- 221, 1.128; 173- 180, 1.105; 182- 188, 1.097; 128- 134, 1.078; 147- 160, 1.185; 431- 469, 1.232; 51-79, 1.227; 195- 205, 1.085; 264- 269, 1.054; 343- 361, 1.222; 42-48, 1.104; 81- 106, 1.165; 365- 379, 1.099; 311- 338, 1.193; 223- 244, 1.245; 277- 286, 1.204;	MYRISTYL 57-62; PKC_PHOSPHO_SITE 480-482; MYRISTYL 26-31; PKC_PHOSPHO_SITE 475-477; CK2_PHOSPHO_SITE 500-503; PKC_PHOSPHO_SITE 468-470; PKC_PHOSPHO_SITE 493-495; CK2_PHOSPHO_SITE 304-307; MYRISTYL 485-490; CK2_PHOSPHO_SITE 310-313; CAMP_PHOSPHO_SITE 505-508; CK2_PHOSPHO_SITE 160-163; CK2_PHOSPHO_SITE 375-378; PKC_PHOSPHO_SITE 38-40; CK2_PHOSPHO_SITE 480-483; CK2_PHOSPHO_SITE 222-225; CK2_PHOSPHO_SITE 282-285; AMIDATION 134-137; CK2_PHOSPHO_SITE 250-253; PKC_PHOSPHO_SITE 282-284; ASN GLYCOSYLATION 115-118; MYRISTYL 86-91; MYRISTYL 451-456; CK2_PHOSPHO_SITE 388-391; CK2_PHOSPHO_SITE 89-92; CK2_PHOSPHO_SITE 21-24; MYRISTYL 392-397; PKC PHOSPHO SITE	CYTOCHROME_C 450-455; GUANYLATE_CYCLASE_2 45-81; ATP_GTP_A 189-196;

			13-20, 1.068; 289- 309, 1.22; 476- 482, 1.052;	360-362;	
DEX0452_016.orf.	N	0 - o1- 2 306;	163- 177, 1.099; 9-42, 1.246; 193- 223, 1.135; 282- 299, 1.158; 141- 159, 1.222; 62-67, 1.054; 109- 136, 1.193; 75-84, 1.204; 274- 280, 1.052; 229- 267, 1.232; 87-107, 1.22;	CK2_PHOSPHO_SITE 186- 189; PKC_PHOSPHO_SITE 266-268; CK2_PHOSPHO_SITE 80-83; PKC_PHOSPHO_SITE 273- 275; MYRISTYL 190-195; CAMP_PHOSPHO_SITE 303- 306; CK2_PHOSPHO_SITE 102-105; PKC_PHOSPHO_SITE 80-82; CK2_PHOSPHO_SITE 108- 111; PKC_PHOSPHO_SITE 291-293; CK2_PHOSPHO_SITE 298- 301; PKC_PHOSPHO_SITE 158-160; CK2_PHOSPHO_SITE 278- 281; MYRISTYL 283-288; CK2_PHOSPHO_SITE 173- 176; CK2_PHOSPHO_SITE 48-51; PKC_PHOSPHO_SITE 278-280; MYRISTYL 249- 254;	CYTOCHROME_C 248-253;
DEX0452_016.aa.2	N	0 - o1- 429;	116- 126, 1.085; 405- 422, 1.158; 144- 165, 1.245; 103- 109, 1.097; 264- 282, 1.222; 68-81, 1.185; 49-55, 1.078; 286- 300, 1.099; 4-27, 1.165; 198- 207, 1.204; 352- 390, 1.232; 316- 346, 1.135; 397- 403, 1.052; 133- 142, 1.128; 210- 230, 1.22; 185- 190, 1.054; 232- 259, 1.193;	CK2_PHOSPHO_SITE 296- 299; CK2_PHOSPHO_SITE 10-13; MYRISTYL 372- 377; CK2_PHOSPHO_SITE 309-312; CK2_PHOSPHO_SITE 231- 234; CK2_PHOSPHO_SITE 225-228; PKC_PHOSPHO_SITE 281- 283; ASN_GLYCOSYLATION 36-39; MYRISTYL 406- 411; PKC_PHOSPHO_SITE 414-416; PKC_PHOSPHO_SITE 401- 403; CK2_PHOSPHO_SITE 401-404; PKC_PHOSPHO_SITE 396- 398; CK2_PHOSPHO_SITE 203-206; AMIDATION 55- 58; CAMP_PHOSPHO_SITE 426-429; MYRISTYL 7-12; PKC_PHOSPHO_SITE 389- 391; CK2_PHOSPHO_SITE 81-84; CK2_PHOSPHO_SITE 421-424; PKC_PHOSPHO_SITE 203- 205; CK2_PHOSPHO_SITE 143-146; MYRISTYL 313- 318; CK2_PHOSPHO_SITE 171-174;	CYTOCHROME_C 371-376; ATP_GTP_A 110- 117;

			94- 101, 1.105;		
DEX0452_016.orf.	N	0 - o1- 306;	75-84, 1.204; 109- 136, 1.193; 141- 159, 1.222; 282- 299, 1.158; 62-67, 1.054; 87-107, 1.22; 229- 267, 1.232; 274- 280, 1.052; 193- 223, 1.135; 163- 177, 1.099; 9-42, 1.246;	CK2_PHOSPHO_SITE 186- 189; CK2_PHOSPHO_SITE 173-176; PKC_PHOSPHO_SITE 278- 280; PKC_PHOSPHO_SITE 266-268; MYRISTYL 249- 254; CK2_PHOSPHO_SITE 278-281; CAMP_PHOSPHO_SITE 303- 306; PKC_PHOSPHO_SITE 80-82; MYRISTYL 190- 195; CK2_PHOSPHO_SITE 48-51; PKC_PHOSPHO_SITE 158-160; CK2_PHOSPHO_SITE 108- 111; PKC_PHOSPHO_SITE 291-293; CK2_PHOSPHO_SITE 80-83; MYRISTYL 283-288; CK2_PHOSPHO_SITE 298- 301; CK2_PHOSPHO_SITE 102-105; PKC_PHOSPHO_SITE 273- 275;	CYTOCHROME_C 248-253;
DEX0452_016.orf.	N	0 - o1- 306;	75-84, 1.204; 274- 280, 1.052; 109- 136, 1.193; 282- 299, 1.158; 87-107, 1.22; 229- 267, 1.232; 163- 177, 1.099; 9-42, 1.246; 62-67, 1.054; 141- 159, 1.222; 193- 223, 1.135;	CK2_PHOSPHO_SITE 102- 105; PKC_PHOSPHO_SITE 158-160; MYRISTYL 190- 195; PKC_PHOSPHO_SITE 80-82; MYRISTYL 283- 288; CK2_PHOSPHO_SITE 48-51; CK2_PHOSPHO_SITE 80-83; CAMP_PHOSPHO_SITE 303- 306; PKC_PHOSPHO_SITE 266-268; MYRISTYL 249- 254; PKC_PHOSPHO_SITE 291-293; CK2_PHOSPHO_SITE 108- 111; CK2_PHOSPHO_SITE 173-176; CK2_PHOSPHO_SITE 298- 301; PKC_PHOSPHO_SITE 273-275; CK2_PHOSPHO_SITE 186- 189; PKC_PHOSPHO_SITE 278-280; CK2_PHOSPHO_SITE 278- 281;	CYTOCHROME_C 248-253;
DEX0452_016.orf.	N	0 - o1- 306;	9-42, 1.246; 109- 136, 1.193; 163- 177, 1.099; 141- 159, 1.222; 87-107, 1.22; 274-	CK2_PHOSPHO_SITE 298- 301; CK2_PHOSPHO_SITE 102-105; MYRISTYL 283- 288; PKC_PHOSPHO_SITE 273-275; MYRISTYL 249- 254; CK2_PHOSPHO_SITE 173-176; PKC_PHOSPHO_SITE 158- 160; CK2_PHOSPHO SITE	CYTOCHROME_C 248-253;

			280, 1.052; 75-84, 1.204; 62-67, 1.054; 229- 267, 1.232; 282- 299, 1.158; 193- 223, 1.135;	48-51; PKC_PHOSPHO_SITE 291-293; CK2_PHOSPHO_SITE 186- 189; PKC_PHOSPHO_SITE 278-280; CAMP_PHOSPHO_SITE 303- 306; MYRISTYL 190-195; CK2_PHOSPHO_SITE 108- 111; PKC_PHOSPHO_SITE 80-82; CK2_PHOSPHO_SITE 278-281; CK2_PHOSPHO_SITE 80-83; PKC_PHOSPHO_SITE 266- 268;	
DEX0452_016.orf.	N	0 - 01- 6	274- 280, 1.052; 62-67, 1.054; 75-84, 1.204; 229- 267, 1.232; 141- 159, 1.222; 87-107, 1.22; 163- 177, 1.099; 109- 136, 1.193; 282- 299, 1.158; 9-42, 1.246; 193- 223, 1.135;	CK2_PHOSPHO_SITE 48-51; PKC_PHOSPHO_SITE 80-82; CK2_PHOSPHO_SITE 108- 111; MYRISTYL 190-195; CK2_PHOSPHO_SITE 80-83; CK2_PHOSPHO_SITE 298- 301; MYRISTYL 249-254; CK2_PHOSPHO_SITE 102- 105; MYRISTYL 283-288; PKC_PHOSPHO_SITE 291- 293; PKC_PHOSPHO_SITE 266-268; CK2_PHOSPHO_SITE 278- 281; CK2_PHOSPHO_SITE 173-176; PKC_PHOSPHO_SITE 158- 160; PKC_PHOSPHO_SITE 278-280; CK2_PHOSPHO_SITE 186- 189; CAMP_PHOSPHO_SITE 303-306; PKC_PHOSPHO_SITE 273- 275;	CYTOCHROME_C 248-253;
DEX0452_017.aa.1	N	0 - 01- 121;	28-49, 1.157; 73-83, 1.115; 88-93, 1.083; 105- 114, 1.146; 61-68, 1.076;	PKC_PHOSPHO_SITE 112- 114; PKC_PHOSPHO_SITE 3-5; MYRISTYL 25-30; PKC_PHOSPHO_SITE 17-19; ASN_GLYCOSYLATION 57- 60; CK2_PHOSPHO_SITE 3- 6; MYRISTYL 33-38; RGD 94-96;	
DEX0452_018.orf.	N	1 - 01- 91; tm 92- 114; i 115- 125;	10-17, 1.079; 57-83, 1.149; 89- 116, 1.205; 38-51, 1.212; 21-30, 1.063;	MYRISTYL 59-64; PKC_PHOSPHO_SITE 83-85; ASN_GLYCOSYLATION 18- 21; PKC_PHOSPHO_SITE 73-75; MYRISTYL 6-11; CAMP_PHOSPHO_SITE 33- 36; MYRISTYL 3-8; CK2_PHOSPHO_SITE 23-26;	
DEX0452_018.aa.1	N	0 - 01- 764;	410- 417, 1.099; 195- 214, 1.168; 369-	MYRISTYL 284-289; CK2_PHOSPHO_SITE 382- 385; MYRISTYL 503-508; CK2_PHOSPHO_SITE 290- 293; CK2_PHOSPHO SITE	

			<p>374,1.091; 446- 454,1.124; 236- 241,1.079; 743- 757,1.218; 654- 665,1.206; 4-9,1.056; 473- 481,1.111; 436- 444,1.104; 554- 587,1.151; 84-96,1.25; 380- 386,1.115; 460- 467,1.061; 22-73,1.243; 110- 134,1.091; 705- 739,1.174; 503- 509,1.136; 247- 253,1.094; 11-20,1.082; 263- 276,1.078; 673- 679,1.051; 99- 108,1.065; 484- 499,1.112; 389- 397,1.136; 295- 303,1.102; 423- 430,1.079; 326- 358,1.218; 589- 650,1.171; 685- 699,1.121; 516- 550,1.148; 309- 317,1.082;</p>	<p>508-511; CK2_PHOSPHO_SITE 374- 377; MYRISTYL 114-119; PKC_PHOSPHO_SITE 333- 335; MYRISTYL 690-695; MYRISTYL 288-293; PKC_PHOSPHO_SITE 436- 438; PKC_PHOSPHO_SITE 680-682; MYRISTYL 709- 714; MYRISTYL 77-82; PKC_PHOSPHO_SITE 540- 542; PKC_PHOSPHO_SITE 514-516; AMIDATION 565- 568; MYRISTYL 20-25; MYRISTYL 499-504; PKC_PHOSPHO_SITE 365- 367; MYRISTYL 617-622; ASN_GLYCOSYLATION 318- 321; PKC_PHOSPHO_SITE 357-359; MYRISTYL 134- 139; ASN_GLYCOSYLATION 697-700; CK2_PHOSPHO_SITE 153- 156; CK2_PHOSPHO_SITE 455-458; PKC_PHOSPHO_SITE 80-82; CK2_PHOSPHO_SITE 57-60; ASN_GLYCOSYLATION 604- 607; MYRISTYL 304-309;</p>	
DEX0452_019.aa.1	N	0 - o1- 267;	<p>159- 166,1.088; 137- 156.1.186:</p>	<p>MYRISTYL 131-136; CK2_PHOSPHO_SITE 234- 237; PKC_PHOSPHO_SITE 251-253: MYRISTYL 146-</p>	

			57-63, 1.129; 240- 246, 1.066; 211- 231, 1.229; 31-38, 1.043; 18-28, 1.075; 68-83, 1.156; 87- 114, 1.168;	151; MYRISTYL 55-60; PKC_PHOSPHO_SITE 189- 191; MYRISTYL 127-132; PKC_PHOSPHO_SITE 123- 125; MYRISTYL 47-52;	
DEX0452_019.orf.	N	0 - o1- 164;	108- 128, 1.229; 6-13, 1.073; 137- 143, 1.066; 34-53, 1.186; 56-63, 1.088;	CK2_PHOSPHO_SITE 131- 134; MYRISTYL 24-29; MYRISTYL 28-33; PKC_PHOSPHO_SITE 86-88; PKC_PHOSPHO_SITE 20-22; PKC_PHOSPHO_SITE 148- 150; MYRISTYL 43-48;	
DEX0452_020.aa.1	N	1 - i1- 33;tm 34- 56;o5 7-99;	75-82, 1.088; 29-46, 1.176; 48-56, 1.119; 59-67, 1.232; 4-17, 1.115;	RGD 75-77; CK2_PHOSPHO_SITE 55-58; MYRISTYL 5-10; CK2_PHOSPHO_SITE 14-17;	SPASE_I_1 66- 73;
DEX0452_020.orf.	N	0 - o1- 136;	52-61, 1.129; 120- 132, 1.15; 63-73, 1.091; 18-43, 1.178; 78- 102, 1.112; 4-14, 1.162;	CK2_PHOSPHO_SITE 65-68; MYRISTYL 63-68; PKC_PHOSPHO_SITE 105- 107; MYRISTYL 36-41; MYRISTYL 84-89; AMIDATION 109-112;	
DEX0452_021.aa.1	N	0 - o1- 139;	121- 136, 1.144; 111- 118, 1.083; 47-60, 1.095; 24-35, 1.136; 17-22, 1.067; 70- 103, 1.137;	CK2_PHOSPHO_SITE 11-14; PKC_PHOSPHO_SITE 4-6; MYRISTYL 61-66; PKC_PHOSPHO_SITE 68-70;	EP450I 69-95; EP450I 112-130; EP450I 49-66; P450 60-77; P450 113-124;
DEX0452_021.orf.	N	0 - o1- 165;	7-12, 1.074; 52-58, 1.095; 63-82, 1.137; 94-101, 1.15; 140- 162, 1.262; 37-45, 1.122; 117- 127, 1.076; 18-28, 1.125; 129- 138, 1.156; 104- 112, 1.114;	ASN GLYCOSYLATION 48- 51; CK2_PHOSPHO_SITE 44-47; MYRISTYL 158- 163; CK2_PHOSPHO_SITE 137-140; CAMP_PHOSPHO_SITE 128- 131; MYRISTYL 8-13; PKC_PHOSPHO_SITE 58-60; MYRISTYL 28-33;	
DEX0452_022.aa.1	N	0 - o1- 136;	21-27, 1.075; 9-18, 1.06; 31-53, 1.153; 82-	MYRISTYL 13-18; PKC_PHOSPHO_SITE 107- 109; PKC_PHOSPHO_SITE 75-77; CK2 PHOSPHO SITE	

			101, 1.124; 128- 133, 1.088; 116- 123, 1.101;	17-20; MYRISTYL 21-26; PKC_PHOSPHO_SITE 28-30; PKC_PHOSPHO_SITE 97-99; CK2_PHOSPHO_SITE 7-10;	
DEX0452_023.aa.1	N	0 - o1- 196;	6-30, 1.222; 160- 168, 1.14; 95- 115, 1.216; 129- 141, 1.073; 33-39, 1.098; 185- 193, 1.186; 74-89, 1.266; 44-56, 1.182;	MYRISTYL 3-8; PKC_PHOSPHO_SITE 144- 146; PKC_PHOSPHO_SITE 179-181; CK2_PHOSPHO_SITE 110- 113; CK2_PHOSPHO_SITE 118-121;	
DEX0452_024.aa.1	N	0 - i1- 69;	14-30, 1.142; 55-66, 1.125; 33-41, 1.09;		
DEX0452_024.orf.1	N	0 - i1- 69;	33-41, 1.09; 55-66, 1.125; 14-30, 1.165;		
DEX0452_025.aa.1	Y	0 - o1- 174;	58-69, 1.244; 30-37, 1.122; 118- 132, 1.119; 81-95, 1.115; 148- 168, 1.153; 98- 104, 1.055; 5-17, 1.131;	MICROBODIES_CTER 172- 174; MYRISTYL 21-26; MYRISTYL 140-145; MYRISTYL 143-148; PKC_PHOSPHO_SITE 24-26; PKC_PHOSPHO_SITE 133- 135; AMIDATION 133-136; CK2_PHOSPHO_SITE 51-54;	
DEX0452_025.orf.1	N	0 - o1- 167;		CK2_PHOSPHO_SITE 44-47; MYRISTYL 136-141; AMIDATION 126-129; PKC_PHOSPHO_SITE 126- 128; MYRISTYL 14-19; MICROBODIES_CTER 165- 167; MYRISTYL 133-138; PKC_PHOSPHO_SITE 17-19;	
DEX0452_026.aa.1	N	0 - i1- 125;	94- 100, 1.057; 6-13, 1.098; 102- 115, 1.134; 21-32, 1.156; 71-77, 1.074;	CK2_PHOSPHO_SITE 33-36; PKC_PHOSPHO_SITE 120- 122; CK2_PHOSPHO_SITE 48-51; ASN GLYCOSYLATION 43- 46; CK2_PHOSPHO_SITE 47-50; PKC_PHOSPHO_SITE 79-81; MYRISTYL 119- 124; PKC_PHOSPHO_SITE 21-23;	TROPOMYOSIN 31- 54; TROPOMYOSIN 87-112;
DEX0452_027.aa.1	N	0 - o1- 106;	60-84, 1.129; 47-52, 1.094; 90- 103, 1.176; 28-44, 1.18;	MYRISTYL 86-91; AMIDATION 14-17; CK2_PHOSPHO_SITE 64-67; CK2_PHOSPHO_SITE 40-43; PKC_PHOSPHO_SITE 4-6;	
DEX0452	N	0 -	60-67. 1.077:	MYRISTYL 203-208:	efhand 59-87:

027.aa.2		o1- 237;	223- 234,1.075; 184- 196,1.196; 148- 169,1.09; 92- 102,1.123; 12-39,1.095; 206- 221,1.11; 72-78,1.06; 49-58,1.078;	PKC_PHOSPHO_SITE 220- 222; MYRISTYL 202-207; CK2_PHOSPHO_SITE 102- 105; MYRISTYL 111-116; ASN_GLYCOSYLATION 86- 89; PKC_PHOSPHO_SITE 196-198; PKC_PHOSPHO_SITE 223- 225; ASN_GLYCOSYLATION 113-116; CK2_PHOSPHO_SITE 76-79; PKC_PHOSPHO_SITE 96-98; MYRISTYL 10-15; PKC_PHOSPHO_SITE 37-39;	EF_HAND 68-80; CALFLAGIN 167- 183; EFh 162- 190; sp_Q94743_SORC_ SCHJA 60-117; EF_HAND_2_1 60- 121; EFh 126- 154; CALFLAGIN 125-143; ehand 96-124; EFh 96- 124; EFh 59-87; EF_HAND_2_2 128-185; ehand 162-190; sp_P12815_PCD6_- MOUSE 131-184; EF_HAND 135- 147; ehand 126-154;
DEX0452_028.aa.1	N	0 - o1- 67;	47-56,1.114; 27-45,1.107;	PKC_PHOSPHO_SITE 60-62; PKC_PHOSPHO_SITE 20-22; CK2_PHOSPHO_SITE 44-47; PKC_PHOSPHO_SITE 26-28; CAMP_PHOSPHO_SITE 17- 20; CK2_PHOSPHO_SITE 4- 7; CK2_PHOSPHO_SITE 20- 23;	
DEX0452_028.orf.1	N	2 - o1- 63;tm 64- 86;i8 7- 92;tm 93- 115;o 116- 156;	53-143,1.35; 27-46,1.107;	CAMP_PHOSPHO_SITE 17- 20; CK2_PHOSPHO_SITE 20-23; PKC_PHOSPHO_SITE 26-28; PKC_PHOSPHO_SITE 20-22; AMIDATION 47-50; CK2_PHOSPHO_SITE 4-7;	PHE_RICH 89- 138;
DEX0452_029.aa.1	N	0 - o1- 829;	782- 791,1.075; 239- 261,1.137; 363- 370,1.091; 640- 648,1.102; 720- 750,1.169; 658- 676,1.129; 705- 717,1.176; 136- 155,1.162; 410- 418,1.08; 562- 571.1.107:	CK2_PHOSPHO_SITE 379- 382; PKC_PHOSPHO_SITE 161-163; MYRISTYL 499- 504; CK2_PHOSPHO_SITE 621-624; CK2_PHOSPHO_SITE 654- 657; PKC_PHOSPHO_SITE 408-410; PKC_PHOSPHO_SITE 79-81; CK2_PHOSPHO_SITE 269- 272; PKC_PHOSPHO_SITE 98-100; PKC_PHOSPHO_SITE 26-28; CK2_PHOSPHO_SITE 191- 194; PKC_PHOSPHO_SITE 2-4; MYRISTYL 155-160; PKC_PHOSPHO_SITE 789- 791; PKC_PHOSPHO_SITE 179-181; CK2_PHOSPHO_SITE 426-	PRO_RICH 552- 688; PRORICH 659-667; PRORICH 626- 632; PRICHEXTENSN 410-422; PRICHEXTENSN 663-688; PRICHEXTENSN 570-587; PRICHEXTENSN 550-566; PRORICH 579- 588;

			329- 354, 1.14; 181- 193, 1.102; 117- 131, 1.108; 398- 405, 1.08; 201- 224, 1.126; 574- 581, 1.06; 372- 396, 1.151; 757- 779, 1.186; 585- 605, 1.093; 265- 271, 1.074; 465- 509, 1.17; 301- 326, 1.203; 164- 176, 1.115; 80-86, 1.107; 684- 693, 1.128; 427- 447, 1.229; 276- 288, 1.128; 514- 525, 1.152; 797- 823, 1.202; 625- 638, 1.129; 50-66, 1.086; 8-26, 1.101; 607- 614, 1.125; 540- 559, 1.148; 695- 703, 1.08; 32-41, 1.153;	429; MYRISTYL 135-140; PKC_PHOSPHO_SITE 222- 224; PKC_PHOSPHO_SITE 653-655; MYRISTYL 236- 241; MYRISTYL 438-443; CK2_PHOSPHO_SITE 807- 810; PKC_PHOSPHO_SITE 363-365; CK2_PHOSPHO_SITE 613- 616; PKC_PHOSPHO_SITE 421-423; PKC_PHOSPHO_SITE 742- 744; CAMP_PHOSPHO_SITE 347-350; ASN_GLYCOSYLATION 531- 534; CK2_PHOSPHO_SITE 226-229; TYR_PHOSPHO_SITE 780- 788; MYRISTYL 274-279; MYRISTYL 166-171; CK2_PHOSPHO_SITE 588- 591; PKC_PHOSPHO_SITE 620-622; CK2_PHOSPHO_SITE 512- 515; CK2_PHOSPHO_SITE 580-583; PKC_PHOSPHO_SITE 237- 239; CAMP_PHOSPHO_SITE 231-234;	
DEX0452_029.orf. 1	N	0 - o1- 443;	365- 372, 1.091; 303- 328, 1.203; 331- 356, 1.14; 52-68, 1.086; 10-28, 1.101; 400- 407, 1.08; 241-	PKC_PHOSPHO_SITE 239- 241; CK2_PHOSPHO_SITE 228-231; PKC_PHOSPHO_SITE 365- 367; CK2_PHOSPHO_SITE 428-431; MYRISTYL 157- 162; PKC_PHOSPHO_SITE 81-83; PKC_PHOSPHO_SITE 181-183; PKC_PHOSPHO_SITE 423- 425; CAMP PHOSPHO SITE	PRICHEXTENSN 409-434; PRICHEXTENSN 217-229; PRICHEXTENSN 353-370; PRICHEXTENSN 307-328; PRICHEXTENSN 333-349;

			263,1.137; 183- 195,1.102; 278- 290,1.128; 374- 398,1.151; 267- 273,1.074; 119- 133,1.108; 166- 178,1.115; 429- 440,1.165; 203- 226,1.126; 412- 420,1.08; 138- 157,1.162; 34-43,1.153; 82-88,1.107;	233-236; MYRISTYL 168-173; PKC_PHOSPHO_SITE 100-102; MYRISTYL 137-142; CAMP_PHOSPHO_SITE 349-352; PKC_PHOSPHO_SITE 28-30; CK2_PHOSPHO_SITE 193-196; PKC_PHOSPHO_SITE 163-165; MYRISTYL 276-281; CK2_PHOSPHO_SITE 271-274; CK2_PHOSPHO_SITE 381-384; MYRISTYL 238-243; PKC_PHOSPHO_SITE 224-226; PKC_PHOSPHO_SITE 4-6; PKC_PHOSPHO_SITE 410-412;	
DEX0452_029.aa.2	Y	0 - o1- 138;	66-88,1.186; 4-14,1.103; 91- 100,1.075; 22-34,1.175; 48-64,1.13; 106- 132,1.202;	MYRISTYL 30-35; CK2_PHOSPHO_SITE 116-119; TYR_PHOSPHO_SITE 89-97; MYRISTYL 37-42; PKC_PHOSPHO_SITE 98-100; MYRISTYL 34-39;	
DEX0452_030.aa.1	N	0 - o1- 60;	20-26,1.085; 30-47,1.11;	MICROBODIES_CTER 58-60; CK2_PHOSPHO_SITE 26-29; CK2_PHOSPHO_SITE 17-20; PKC_PHOSPHO_SITE 49-51;	
DEX0452_030.orf.1	Y	2 - i1- 19;tm 20- 39;o4 0- 58;tm 59- 81;i8 2-99;	32-38,1.081; 12-28,1.203; 56-95,1.222; 4-10,1.135;	MICROBODIES_CTER 97-99; MYRISTYL 38-43;	
DEX0452_031.aa.1	Y	0 - o1- 294;	127- 143,1.165; 222- 228,1.056; 166- 187,1.178; 230- 236,1.035; 76-86,1.208; 256- 269,1.186; 63-70,1.076; 101-	CK2_PHOSPHO_SITE 197-200; CK2_PHOSPHO_SITE 233-236; MYRISTYL 33-38; CK2_PHOSPHO_SITE 170-173; CK2_PHOSPHO_SITE 147-150; PKC_PHOSPHO_SITE 197-199; MYRISTYL 178-183; CK2_PHOSPHO_SITE 51-54; MYRISTYL 229-234; PKC_PHOSPHO_SITE 220-222; MYRISTYL 269-274; CK2_PHOSPHO SITE	complex1_24kD 53-209; COMPLEX1_24K 166-184; sp_Q9BV41_Q9BV41_HUMAN 58-138;

			119,1.19; 36-43,1.092; 148- 159,1.071; 4-15,1.064; 242- 253,1.236; 276- 291,1.19;	163-166; MYRISTYL 157- 162;	
DEX0452_031.aa.2	Y	0 - o1- 250;	4-15,1.064; 36-43,1.092; 166- 187,1.178; 76-86,1.208; 222- 228,1.056; 101- 119,1.19; 148- 159,1.071; 127- 143,1.165; 63-70,1.076;	CK2_PHOSPHO_SITE 170- 173; CK2_PHOSPHO_SITE 233-236; PKC_PHOSPHO_SITE 197- 199; MYRISTYL 178-183; PKC_PHOSPHO_SITE 235- 237; MICROBODIES_CTER 248-250; MYRISTYL 229- 234; CK2_PHOSPHO_SITE 147-150; CK2_PHOSPHO_SITE 197- 200; CK2_PHOSPHO_SITE 163-166; MYRISTYL 157- 162; MYRISTYL 33-38; PKC_PHOSPHO_SITE 220- 222; CK2_PHOSPHO_SITE 51-54;	complex1_24kD 53-209; sp_Q9BV41_Q9BV4 1_HUMAN 58-138; COMPLEX1_24K 166-184;
DEX0452_031.aa.3	Y	0 - o1- 232;	76-86,1.208; 161- 229,1.255; 36-43,1.092; 4-15,1.064; 63-70,1.076; 127- 143,1.165; 148- 155,1.071; 101- 119,1.19;	CK2_PHOSPHO_SITE 51-54; PKC_PHOSPHO_SITE 219- 221; CK2_PHOSPHO_SITE 147-150; MYRISTYL 33- 38;	sp_Q9BV41_Q9BV4 1_HUMAN 58-138; complex1_24kD 53-194;
DEX0452_032.aa.1	N	0 - o1- 28;	5-23,1.13;		
DEX0452_032.orf.1	N	0 - o1- 106;	17-53,1.155; 63-96,1.143;	ASN_GLYCOSYLATION 16- 19; CK2_PHOSPHO_SITE 70-73; MYRISTYL 55-60; MYRISTYL 6-11; PKC_PHOSPHO_SITE 58-60; MYRISTYL 82-87; MYRISTYL 40-45;	
DEX0452_033.aa.1	Y	0 - o1- 137;	63-85,1.083; 109- 120,1.158; 51-61,1.086; 4-36,1.212; 90-98,1.121;	PKC_PHOSPHO_SITE 53-55; MYRISTYL 132-137; CK2_PHOSPHO_SITE 85-88; CAMP_PHOSPHO_SITE 55- 58; ASN_GLYCOSYLATION 104-107; MYRISTYL 116- 121;	
DEX0452_033.aa.2	Y	0 - o1- 241;	51-61,1.086; 109- 120,1.158;	ASN_GLYCOSYLATION 206- 209; PKC_PHOSPHO_SITE 134-136;	

			139- 145, 1.045; 154- 167, 1.085; 4-36, 1.212; 208- 238, 1.237; 63-85, 1.083; 177- 190, 1.044; 90-98, 1.121; 125- 131, 1.066;	CK2_PHOSPHO_SITE 210- 213; CK2_PHOSPHO_SITE 85-88; ASN_GLYCOSYLATION 104- 107; CK2_PHOSPHO_SITE 178-181; CAMP_PHOSPHO_SITE 55- 58; PKC_PHOSPHO_SITE 53-55; ASN_GLYCOSYLATION 174- 177; MYRISTYL 116-121;	
DEX0452_034.aa.1	N	0 - o1- 102;		AMIDATION 63-66; CAMP_PHOSPHO_SITE 93- 96; PKC_PHOSPHO_SITE 92-94; MYRISTYL 34-39; CK2_PHOSPHO_SITE 16-19; PKC_PHOSPHO_SITE 16-18; MYRISTYL 46-51;	
DEX0452_034.aa.2	N	0 - o1- 207;	98-107, 1.08; 161- 168, 1.133; 177- 186, 1.14; 44-51, 1.08; 113- 141, 1.174; 62-82, 1.114; 194- 204, 1.117;	ASN_GLYCOSYLATION 58- 61; CK2_PHOSPHO_SITE 26-29; MYRISTYL 35-40; CK2_PHOSPHO_SITE 79-82; PKC_PHOSPHO_SITE 52-54; CK2_PHOSPHO_SITE 167- 170; CK2_PHOSPHO_SITE 129-132; MYRISTYL 151- 156; CK2_PHOSPHO_SITE 16-19; MYRISTYL 44-49;	
DEX0452_034.orf.2	N	0 - o1- 208;	62-81, 1.179; 178- 187, 1.14; 39-45, 1.068; 99-108, 1.08; 114- 142, 1.174; 195- 205, 1.117; 29-37, 1.075; 162- 169, 1.133;	MYRISTYL 9-14; PKC_PHOSPHO_SITE 18-20; AMIDATION 23-26; CAMP_PHOSPHO_SITE 47- 50; CK2_PHOSPHO_SITE 168-171; PKC_PHOSPHO_SITE 46-48; MYRISTYL 152-157; CK2_PHOSPHO_SITE 50-53; MYRISTYL 11-16; CK2_PHOSPHO_SITE 130- 133;	PRO_RICH 31- 111;
DEX0452_035.aa.1	N	0 - o1- 267;	258- 264, 1.107; 187- 212, 1.258; 42-72, 1.128; 82- 100, 1.189; 143- 173, 1.178; 177- 184, 1.121; 219- 235, 1.103; 11-27, 1.136;	CK2_PHOSPHO_SITE 160- 163; MYRISTYL 75-80; MYRISTYL 90-95; CK2_PHOSPHO_SITE 235- 238; PKC_PHOSPHO_SITE 177-179; MYRISTYL 95- 100; MYRISTYL 79-84; MYRISTYL 115-120; MYRISTYL 76-81; CK2_PHOSPHO_SITE 197- 200; MYRISTYL 173-178;	UBIQUITIN_2 176-267;
DEX0452	Y	0 -	177-	CK2 PHOSPHO SITE 118-	UBIQUITIN 2

035.orf. 1		o1- 225;	193,1.103; 216- 222,1.107; 4-27,1.199; 135- 142,1.121; 145- 170,1.258; 33-54,1.128; 101- 131,1.178;	121; MYRISTYL 8-13; PKC_PHOSPHO_SITE 135- 137; CAMP_PHOSPHO_SITE 27-30; CK2_PHOSPHO_SITE 193-196; MYRISTYL 73- 78; MYRISTYL 131-136; PKC_PHOSPHO_SITE 1-3; CK2_PHOSPHO_SITE 155- 158;	134-225;
DEX0452_ 036.aa.1	N	0 - o1- 224;	75-96,1.128; 143- 173,1.178; 42-69,1.128; 177- 184,1.121; 11-27,1.136; 187- 212,1.258;	MYRISTYL 115-120; CK2_PHOSPHO_SITE 160- 163; CK2_PHOSPHO_SITE 197-200; CAMP_PHOSPHO_SITE 69- 72; PKC_PHOSPHO_SITE 177-179; MYRISTYL 173- 178;	
DEX0452_ 036.aa.2	N	143- 173,1.178; 188- 196,1.112; 177- 184,1.121; 246;t m247- 11-27,1.136; 269;i 270- 300;	143- 173,1.178; 188- 196,1.112; 177- 184,1.121; 246;t m247- 11-27,1.136; 269;i 270- 300;	MYRISTYL 173-178; CAMP_PHOSPHO_SITE 69- 72; MYRISTYL 115-120; MYRISTYL 227-232; CK2_PHOSPHO_SITE 273- 276; MYRISTYL 225-230; CK2_PHOSPHO_SITE 204- 207; PKC_PHOSPHO_SITE 177-179; CK2_PHOSPHO_SITE 160- 163;	
DEX0452_ 036.orf.	N	0 - o1- 236;	9-22,1.082; 103- 124,1.128; 171- 201,1.178; 205- 212,1.121; 39-55,1.136; 216- 230,1.117; 70-97,1.128;	CAMP_PHOSPHO_SITE 97- 100; MYRISTYL 24-29; CK2_PHOSPHO_SITE 188- 191; MYRISTYL 201-206; PKC_PHOSPHO_SITE 205- 207; ASN GLYCOSYLATION 26-29; MYRISTYL 143- 148; PKC_PHOSPHO_SITE 233-235;	
DEX0452_ 037.aa.1	N	0 - o1- 143;	120- 134,1.162; 72-90,1.076; 37-63,1.15; 92- 100,1.089; 16-25,1.138;	PKC_PHOSPHO_SITE 10-12; CK2_PHOSPHO_SITE 76-79; PKC_PHOSPHO_SITE 63-65; MYRISTYL 120-125; CK2_PHOSPHO_SITE 10-13; CK2_PHOSPHO_SITE 132- 135; CK2_PHOSPHO_SITE 31-34; MYRISTYL 103- 108;	
DEX0452_ 037.orf.	N	0 - o1- 126;	21-32,1.126; 64-73,1.138; 85-111,1.15; 40-49,1.093; 4-17,1.09;	PKC_PHOSPHO_SITE 58-60; PKC_PHOSPHO_SITE 111- 113; CK2_PHOSPHO_SITE 79-82; PKC_PHOSPHO_SITE 11-13; CK2_PHOSPHO SITE	

				58-61; MYRISTYL 8-13; PKC_PHOSPHO_SITE 75-77; PKC_PHOSPHO_SITE 40-42; RGD 45-47; MYRISTYL 11-16; MYRISTYL 91-96; MYRISTYL 68-73; PKC_PHOSPHO_SITE 32-34; CK2_PHOSPHO_SITE 40-43; MYRISTYL 81-86; PKC_PHOSPHO_SITE 82-84;	
DEX0452_037.aa.2	N	0 - o1- 116;	57-69, 1.097; 29-39, 1.075; 95- 113, 1.139; 86-92, 1.093; 13-24, 1.192;	MYRISTYL 59-64; PKC_PHOSPHO_SITE 4-6; CAMP_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 9-11; CK2_PHOSPHO_SITE 73-76; PKC_PHOSPHO_SITE 45-47; MYRISTYL 67-72;	
DEX0452_038.aa.1	N	0 - o1- 77;	11-54, 1.18; 60-74, 1.178;	PKC_PHOSPHO_SITE 25-27; PHE_RICH 3-89;	
DEX0452_038.orf.1	Y	3 - o1- 27;tm 28- 50;i5 1- 54;tm 55- 74;o7 5- 88;tm 89- 111;i 112- 115;	12-75, 1.243; 78-84, 1.073; 87- 112, 1.186;		
DEX0452_038.orf.2	N	0 - o1- 84;		PKC_PHOSPHO_SITE 9-11; MYRISTYL 59-64; ASN_GLYCOSYLATION 79-82; MYRISTYL 77-82; CAMP_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 45-47; PKC_PHOSPHO_SITE 4-6; MYRISTYL 67-72; CK2_PHOSPHO_SITE 73-76;	
DEX0452_038.orf.3	N	0 - o1- 84;		ASN_GLYCOSYLATION 79-82; MYRISTYL 67-72; PKC_PHOSPHO_SITE 4-6; CK2_PHOSPHO_SITE 73-76; CAMP_PHOSPHO_SITE 6-9; PKC_PHOSPHO_SITE 9-11; MYRISTYL 59-64; MYRISTYL 77-82; PKC_PHOSPHO_SITE 45-47;	
DEX0452_039.aa.1	N	0 - o1- 104;	51-59, 1.155; 69-74, 1.044; 4-30, 1.153; 36-42, 1.041;	PKC_PHOSPHO_SITE 73-75; PKC_PHOSPHO_SITE 41-43; CK2_PHOSPHO_SITE 73-76; MYRISTYL 51-56; MYRISTYL 28-33; PKC_PHOSPHO_SITE 59-61; PKC_PHOSPHO SITE 77-79;	

				TYR_PHOSPHO_SITE 42-48; ASN GLYCOSYLATION 35-38; MYRISTYL 91-96; CK2_PHOSPHO_SITE 77-80; PKC_PHOSPHO_SITE 40-42; MYRISTYL 12-17;	
DEX0452_039.orf. 1	N	0 - o1- 107;	18-25, 1.065; 31-37, 1.041; 64-69, 1.044; 46-54, 1.155; 6-16, 1.153; 92-99, 1.114;	MYRISTYL 86-91; PKC_PHOSPHO_SITE 3-5; PKC_PHOSPHO_SITE 35-37; TYR_PHOSPHO_SITE 37-43; ASN GLYCOSYLATION 30-33; CK2_PHOSPHO_SITE 68-71; ASN GLYCOSYLATION 1-4; PKC_PHOSPHO_SITE 36-38; MYRISTYL 23-28; MYRISTYL 46-51; CK2_PHOSPHO_SITE 102-105; PKC_PHOSPHO_SITE 68-70; CK2_PHOSPHO_SITE 72-75; PKC_PHOSPHO_SITE 54-56; PKC_PHOSPHO_SITE 72-74;	
DEX0452_040_aa.1	N	0 - o1- 38;		MICROBODIES_CTER 36-38; MYRISTYL 25-30;	
DEX0452_040.orf. 1	N	0 - o1- 47;	32-44, 1.181;	CK2_PHOSPHO_SITE 13-16; PKC_PHOSPHO_SITE 15-17; CK2_PHOSPHO_SITE 34-37; CAMP_PHOSPHO_SITE 17-20; PKC_PHOSPHO_SITE 20-22; ASN GLYCOSYLATION 11-14;	
DEX0452_041_aa.1	N	0 - o1- 71;	22-46, 1.114; 15-20, 1.017;	MYRISTYL 53-58; PKC_PHOSPHO_SITE 32-34; CK2_PHOSPHO_SITE 48-51; PKC_PHOSPHO_SITE 36-38; CAMP_PHOSPHO_SITE 33-36; CK2_PHOSPHO_SITE 44-47;	
DEX0452_042_aa.1	N	0 - o1- 138;	116- 135, 1.156; 5-15, 1.163; 39-46, 1.103; 81-97, 1.184; 100- 110, 1.117; 56-73, 1.221;	CK2_PHOSPHO_SITE 7-10; PKC_PHOSPHO_SITE 33-35; MYRISTYL 17-22; MYRISTYL 63-68; PKC_PHOSPHO_SITE 53-55; PKC_PHOSPHO_SITE 30-32; CAMP_PHOSPHO_SITE 27-30; MYRISTYL 101-106; CK2_PHOSPHO_SITE 33-36; PKC_PHOSPHO_SITE 41-43; CK2_PHOSPHO_SITE 44-47; ZF_FYVE 36-92; FYVE 31-91; FYVE 28-93;	
DEX0452_043_aa.1	N	0 - o1- 67;	4-23, 1.122; 49-55, 1.095;	AMIDATION 25-28; CK2_PHOSPHO_SITE 31-34; PKC_PHOSPHO_SITE 39-41; PKC_PHOSPHO_SITE 58-60; MYRISTYL 25-30; PKC_PHOSPHO SITE 31-33;	

DEX0452_043.orf.	N	0 - o1- 65;		PKC_PHOSPHO_SITE 11-13; PKC_PHOSPHO_SITE 37-39; MYRISTYL 23-28; PKC_PHOSPHO_SITE 29-31; CK2_PHOSPHO_SITE 29-32; PKC_PHOSPHO_SITE 56-58; AMIDATION 23-26;	
DEX0452_043_aa.2	N	0 - o1- 195;	29-81, 1.173; 84- 103, 1.156; 19-27, 1.069; 132- 141, 1.155; 4-14, 1.107; 105- 121, 1.099;	CK2_PHOSPHO_SITE 7-10; MYRISTYL 177-182; PKC_PHOSPHO_SITE 83-85; CK2_PHOSPHO_SITE 129- 132; MYRISTYL 183-188; PKC_PHOSPHO_SITE 102- 104; PKC_PHOSPHO_SITE 170-172;	
DEX0452_044.orf.	N	0 - o1- 124;	83-90, 1.086; 110- 121, 1.125; 24-46, 1.107;	CK2_PHOSPHO_SITE 61-64; PKC_PHOSPHO_SITE 49-51; CAMP_PHOSPHO_SITE 117- 120; PKC_PHOSPHO_SITE 120-122; PKC_PHOSPHO_SITE 18-20; PKC_PHOSPHO_SITE 33-35; CAMP_PHOSPHO_SITE 103- 106;	
DEX0452_044_aa.1	N	0 - o1- 106;	6-28, 1.107; 92- 103, 1.125; 65-72, 1.086;	CAMP_PHOSPHO_SITE 99- 102; CK2_PHOSPHO_SITE 43-46; PKC_PHOSPHO_SITE 31-33; PKC_PHOSPHO_SITE 15-17; CAMP_PHOSPHO_SITE 85- 88; PKC_PHOSPHO_SITE 102-104;	
DEX0452_044.orf.	N	0 - o1- 129;	89-96, 1.086; 111- 117, 1.061; 121- 126, 1.048; 30-52, 1.107;	MYRISTYL 11-16; CK2_PHOSPHO_SITE 67-70; PKC_PHOSPHO_SITE 24-26; ASN GLYCOSYLATION 121- 124; PKC_PHOSPHO_SITE 39-41; PKC_PHOSPHO_SITE 55-57; CAMP_PHOSPHO_SITE 109- 112; CK2_PHOSPHO_SITE 7-10;	
DEX0452_044_aa.2	N	0 - o1- 130;	112- 118, 1.061; 90-97, 1.086; 31-53, 1.107; 10-17, 1.075;	PKC_PHOSPHO_SITE 56-58; CK2_PHOSPHO_SITE 7-10; PKC_PHOSPHO_SITE 25-27; CK2_PHOSPHO_SITE 68-71; ASN GLYCOSYLATION 122- 125; CAMP_PHOSPHO_SITE 110-113; PKC_PHOSPHO_SITE 40-42;	
DEX0452_045.orf.	N	0 - o1- 85;	33-74, 1.28; 4-29, 1.192;	TYR_PHOSPHO_SITE 33-41; CK2_PHOSPHO_SITE 81-84; TYR_PHOSPHO_SITE 35-42;	
DEX0452_045_aa.1	N	0 - o1- 68;	20-27, 1.075; 38-43, 1.04; 9-18, 1.094; 45-52, 1.117;	LEUCINE_ZIPPER 25-46;	

DEX0452_046.orf.	N	0 - o1- 1 378;	83-92, 1.09; 117- 125, 1.175; 227- 238, 1.116; 62-73, 1.161; 196- 208, 1.135; 172- 185, 1.155; 272- 279, 1.112; 4-12, 1.069; 22-29, 1.152; 260- 270, 1.079; 318- 352, 1.157; 145- 156, 1.136;	MYRISTYL 14-19; CK2_PHOSPHO_SITE 253- 256; MYRISTYL 335-340; ASN_GLYCOSYLATION 21- 24; CK2_PHOSPHO_SITE 155-158; CK2_PHOSPHO_SITE 268- 271; ASN_GLYCOSYLATION 282-285; CK2_PHOSPHO_SITE 284- 287; CK2_PHOSPHO_SITE 226-229; MYRISTYL 174- 179; PKC_PHOSPHO_SITE 296-298; CK2_PHOSPHO_SITE 339- 342; CK2_PHOSPHO_SITE 312-315; PKC_PHOSPHO_SITE 33-35; CK2_PHOSPHO_SITE 303- 306; TYR_PHOSPHO_SITE 110-118;	
DEX0452_046_aa.1	N	0 - o1- 1 876;	466- 476, 1.079; 817- 826, 1.106; 402- 414, 1.135; 323- 331, 1.175; 158- 168, 1.066; 95- 101, 1.051; 175- 193, 1.182; 833-873, 1.1; 378- 391, 1.155; 722- 784, 1.123; 228- 235, 1.152; 524- 558, 1.157; 289- 298, 1.09; 18-30, 1.085; 268- 279, 1.161; 71-87, 1.183; 575- 581, 1.072; 663- 709, 1.134; 647- 656, 1.098; 41-59, 1.13; 196- 218, 1.184;	CK2_PHOSPHO_SITE 432- 435; ASN_GLYCOSYLATION 488-491; CK2_PHOSPHO_SITE 490- 493; CK2_PHOSPHO_SITE 459-462; MYRISTYL 145- 150; CK2_PHOSPHO_SITE 474-477; AMIDATION 791- 794; MYRISTYL 220-225; PKC_PHOSPHO_SITE 4-6; CAMP_PHOSPHO_SITE 798- 801; MYRISTYL 380-385; CK2_PHOSPHO_SITE 754- 757; CAMP_PHOSPHO_SITE 5-8; CK2_PHOSPHO_SITE 518-521; CK2_PHOSPHO_SITE 361- 364; CK2_PHOSPHO_SITE 509-512; MYRISTYL 76- 81; PKC_PHOSPHO_SITE 68-70; PKC_PHOSPHO_SITE 34-36; CK2_PHOSPHO_SITE 180-183; CK2_PHOSPHO_SITE 545- 548; PKC_PHOSPHO_SITE 502-504; TYR_PHOSPHO_SITE 316- 324; AMIDATION 796-799; PKC_PHOSPHO_SITE 800- 802; MYRISTYL 541-546; PKC_PHOSPHO_SITE 801- 803; CK2_PHOSPHO_SITE 696-699; PKC_PHOSPHO_SITE 239- 241; ASN_GLYCOSYLATION 227-230; MYRISTYL 148- 153; PKC_PHOSPHO_SITE 10-12; MYRISTYL 39-44;	

			478- 485, 1.112; 588- 625, 1.092; 348- 362, 1.136; 433- 444, 1.116; 130- 141, 1.208;	AMIDATION 807-810;	
DEX0452_046.orf.	N	0 - o1- 378;	62-73, 1.161; 4-12, 1.069; 117- 125, 1.175; 272- 279, 1.112; 172- 185, 1.155; 196- 208, 1.135; 22-29, 1.152; 145- 156, 1.136; 83-92, 1.09; 318- 352, 1.157; 260- 270, 1.079; 227- 238, 1.116;	ASN_GLYCOSYLATION 21- 24; MYRISTYL 14-19; CK2_PHOSPHO_SITE 268- 271; MYRISTYL 174-179; MYRISTYL 335-340; CK2_PHOSPHO_SITE 226- 229; TYR_PHOSPHO_SITE 110-118; CK2_PHOSPHO_SITE 303- 306; CK2_PHOSPHO_SITE 155-158; PKC_PHOSPHO_SITE 296- 298; CK2_PHOSPHO_SITE 312-315; CK2_PHOSPHO_SITE 253- 256; PKC_PHOSPHO_SITE 33-35; CK2_PHOSPHO_SITE 284-287; ASN_GLYCOSYLATION 282- 285; CK2_PHOSPHO_SITE 339-342;	
DEX0452_046.aa.2	N	0 - o1- 782;	81-99, 1.182; 481- 487, 1.072; 102- 124, 1.184; 64-74, 1.066; 36-47, 1.208; 229- 237, 1.175; 384- 391, 1.112; 308- 320, 1.135; 195- 204, 1.09; 339- 350, 1.116; 430- 464, 1.157; 174- 185, 1.161; 553- 562, 1.098; 739-779, 1.1; 723- 732, 1.106; 569- 615, 1.134;	PKC_PHOSPHO_SITE 707- 709; PKC_PHOSPHO_SITE 408-410; CK2_PHOSPHO_SITE 424- 427; CK2_PHOSPHO_SITE 338-341; CK2_PHOSPHO_SITE 602- 605; PKC_PHOSPHO_SITE 145-147; ASN_GLYCOSYLATION 394- 397; TYR_PHOSPHO_SITE 222-230; CK2_PHOSPHO_SITE 660- 663; CK2_PHOSPHO_SITE 380-383; CK2_PHOSPHO_SITE 415- 418; AMIDATION 713-716; MYRISTYL 51-56; ASN_GLYCOSYLATION 133- 136; AMIDATION 702-705; CK2_PHOSPHO_SITE 267- 270; CK2_PHOSPHO_SITE 451-454; CAMP_PHOSPHO_SITE 704- 707; MYRISTYL 447-452; MYRISTYL 54-59; MYRISTYL 286-291; CK2 PHOSPHO SITE 86-89;	

			494- 531,1.092; 372- 382,1.079; 284- 297,1.155; 628- 690,1.123; 254- 268,1.136; 134- 141,1.152;	MYRISTYL 126-131; PKC_PHOSPHO_SITE 706- 708; CK2_PHOSPHO_SITE 365-368; AMIDATION 697- 700; CK2_PHOSPHO_SITE 396-399;	
DEX0452_047.aa.1	N	0 - 01- 449;	173- 179,1.04; 129- 136,1.092; 107- 122,1.23; 431- 444,1.172; 346- 378,1.206; 421- 428,1.175; 302- 317,1.111; 143- 163,1.136; 4-27,1.109; 268- 290,1.163; 184- 215,1.243; 402- 408,1.098; 71- 105,1.238; 249- 259,1.145; 59-68,1.117; 334- 344,1.083; 219- 246,1.153; 31-50,1.154; 324- 332,1.141;	AMIDATION 181-184; TYR_PHOSPHO_SITE 407- 414; CK2_PHOSPHO_SITE 271-274; MYRISTYL 259- 264; PKC_PHOSPHO_SITE 322-324; MYRISTYL 121- 126; CK2_PHOSPHO_SITE 378-381; MYRISTYL 122- 127; PKC_PHOSPHO_SITE 263-265; MYRISTYL 126- 131; CK2_PHOSPHO_SITE 394-397; MYRISTYL 21- 26; MYRISTYL 390-395; CK2_PHOSPHO_SITE 445- 448; PKC_PHOSPHO_SITE 275-277; AMIDATION 17- 20;	ATP_GTP_A 259- 266; Thymidylate_kin 257-438;
DEX0452_048.orf.	N	0 - 01- 1 590;	153- 158,1.105; 31-43,1.115; 571- 582,1.076; 61-72,1.122; 96- 111,1.191; 478- 484,1.063; 556- 563,1.108;	CK2_PHOSPHO_SITE 551- 554; ASN,GLYCOSYLATION 271-274; ASN,GLYCOSYLATION 362- 365; PKC_PHOSPHO_SITE 273-275; CK2_PHOSPHO_SITE 75-78; PKC_PHOSPHO_SITE 11-13; CAMP_PHOSPHO_SITE 211- 214; MYRISTYL 305-310; PKC_PHOSPHO_SITE 187- 189; ASN GLYCOSYLATION	

			<p>236- 242,1.09; 503- 515,1.206; 493- 498,1.061; 127- 134,1.141; 438- 449,1.088; 114- 119,1.031; 248- 254,1.092; 464- 476,1.08; 13-20,1.096; 283- 291,1.065; 80-89,1.135; 414- 427,1.096; 46-52,1.055; 370- 377,1.087; 190- 199,1.138; 166- 175,1.088; 141- 147,1.106; 354- 360,1.024; 401- 407,1.103; 260- 268,1.16; 297- 327,1.109; 525- 549,1.249;</p> <p>449-452; ASN_GLYCOSYLATION 57- 60; PKC_PHOSPHO_SITE 219-221; PKC_PHOSPHO_SITE 44-46; PKC_PHOSPHO_SITE 60-62; ASN_GLYCOSYLATION 416- 419; CK2_PHOSPHO_SITE 412-415; PKC_PHOSPHO_SITE 397- 399; CK2_PHOSPHO_SITE 215-218; ASN_GLYCOSYLATION 463- 466; PKC_PHOSPHO_SITE 24-26; CAMP_PHOSPHO_SITE 275- 278; PKC_PHOSPHO_SITE 121-123; PKC_PHOSPHO_SITE 365- 367; AMIDATION 91-94; CAMP_PHOSPHO_SITE 188- 191; CK2_PHOSPHO_SITE 471-474; CK2_PHOSPHO_SITE 177- 180; CK2_PHOSPHO_SITE 44-47; PKC_PHOSPHO_SITE 89-91; PKC_PHOSPHO_SITE 120-122; ASN_GLYCOSYLATION 162- 165; ASN_GLYCOSYLATION 336-339; CK2_PHOSPHO_SITE 407- 410; CK2_PHOSPHO_SITE 219-222; PKC_PHOSPHO_SITE 460- 462; CK2_PHOSPHO_SITE 518-521; PKC_PHOSPHO_SITE 56-58; CK2_PHOSPHO_SITE 233- 236; ASN_GLYCOSYLATION 400-403; PKC_PHOSPHO_SITE 451- 453; CAMP_PHOSPHO_SITE 586-589; CK2_PHOSPHO_SITE 245- 248; MYRISTYL 504-509; ASN_GLYCOSYLATION 279- 282; MYRISTYL 512-517; PKC_PHOSPHO_SITE 257- 259; PKC_PHOSPHO_SITE 569-571; CK2_PHOSPHO_SITE 121- 124; MYRISTYL 32-37; CK2_PHOSPHO_SITE 172- 175; PKC_PHOSPHO_SITE 301-303; PKC_PHOSPHO_SITE 113- 115; MYRISTYL 269-274; PKC_PHOSPHO SITE 21-23;</p>
--	--	--	--

			CK2_PHOSPHO_SITE 68-71; MYRISTYL 196-201; CK2_PHOSPHO_SITE 362- 365; CK2_PHOSPHO_SITE 581-584; PKC_PHOSPHO_SITE 342- 344; CK2_PHOSPHO_SITE 12-15; CK2_PHOSPHO_SITE 623-626; CK2_PHOSPHO_SITE 106- 109; ASN_GLYCOSYLATION 291-294; CAMP_PHOSPHO_SITE 79- 82; CK2_PHOSPHO_SITE 136-139; ASN_GLYCOSYLATION 170- 173; MYRISTYL 619-624; PKC_PHOSPHO_SITE 351- 353; MYRISTYL 395-400; CAMP_PHOSPHO_SITE 102- 105; CK2_PHOSPHO_SITE 442-445; PKC_PHOSPHO_SITE 4-6; CK2_PHOSPHO_SITE 409- 412; CK2_PHOSPHO_SITE 63-66; CK2_PHOSPHO_SITE 564-567; CAMP_PHOSPHO_SITE 166- 169; PKC_PHOSPHO_SITE 460-462; ASN_GLYCOSYLATION 53- 56; PKC_PHOSPHO_SITE 192-194; ASN_GLYCOSYLATION 354- 357; CK2_PHOSPHO_SITE 124-127; PKC_PHOSPHO_SITE 256- 258; CK2_PHOSPHO_SITE 110-113; ASN_GLYCOSYLATION 307- 310; PKC_PHOSPHO_SITE 11-13; CK2_PHOSPHO_SITE 303-306; MYRISTYL 403- 408; PKC_PHOSPHO_SITE 148-150; ASN_GLYCOSYLATION 162- 165; MYRISTYL 541-546; PKC_PHOSPHO_SITE 12-14; PKC_PHOSPHO_SITE 78-80; CK2_PHOSPHO_SITE 298- 301; ASN_GLYCOSYLATION 253-256; PKC_PHOSPHO_SITE 110- 112; ASN_GLYCOSYLATION 227-230; ASN_GLYCOSYLATION 340- 343; PKC_PHOSPHO_SITE 164-166; PKC PHOSPHO SITE 288-
DEX0452_N	0 -	01- 661;	549- 563, 1.095; 245- 251, 1.024; 81-90, 1.138; 516- 534, 1.113; 608- 615, 1.081; 369- 375, 1.063; 394- 406, 1.206; 139- 145, 1.092; 329- 340, 1.088; 151- 159, 1.16; 416- 440, 1.249; 57-66, 1.088; 127- 133, 1.09; 261- 268, 1.087; 32-38, 1.106; 305- 318, 1.096; 5-10, 1.031; 488-503, 1.1; 188- 218, 1.109; 536- 547, 1.217; 447- 454, 1.108; 575- 596, 1.136; 462- 470, 1.076; 44-49, 1.105; 639- 654, 1.129; 174- 182, 1.065; 292- 298, 1.103; 617-625, 1.1; 355- 367, 1.08; 18-25, 1.141; 384- 389, 1.061;

				290; MYRISTYL 160-165;	
DEX0452_049.aa.1	N	0 - o1- 157;	96- 113, 1.058; 137- 154, 1.127; 51-59, 1.122; 14-34, 1.143; 65-83, 1.119; 40-49, 1.152; 125- 131, 1.094;	MYRISTYL 26-31; MYRISTYL 82-87; CK2_PHOSPHO_SITE 90-93; MYRISTYL 39-44; MYRISTYL 137-142; MYRISTYL 7-12; PKC_PHOSPHO_SITE 90-92; CK2_PHOSPHO_SITE 114- 117; CK2_PHOSPHO_SITE 50-53; MYRISTYL 9-14; MYRISTYL 22-27; MYRISTYL 38-43; CK2_PHOSPHO_SITE 2-5; MYRISTYL 34-39; MYRISTYL 121-126;	
DEX0452_050.aa.1	N	0 - o1- 269;	201- 208, 1.058; 98- 104, 1.102; 115- 121, 1.053; 5-18, 1.056; 159- 170, 1.084; 123- 129, 1.072; 249- 258, 1.062; 225- 244, 1.131; 260- 266, 1.063;	CK2_PHOSPHO_SITE 21-24; CK2_PHOSPHO_SITE 156- 159; MYRISTYL 6-11; TYR_PHOSPHO_SITE 33-40; CK2_PHOSPHO_SITE 54-57; CK2_PHOSPHO_SITE 43-46; LEUCINE_ZIPPER 210-231; MYRISTYL 87-92; MYRISTYL 253-258; CK2_PHOSPHO_SITE 163- 166; CK2_PHOSPHO_SITE 13-16; TYR_PHOSPHO_SITE 22-29; TYR_PHOSPHO_SITE 121-127; PKC_PHOSPHO_SITE 171- 173; PKC_PHOSPHO_SITE 156-158; CK2_PHOSPHO_SITE 192- 195;	EFh 146-174; EFh 94-122; GLU_RICH 24- 223; EF_HAND 155-167; EF_HAND 103- 115; EF_HAND_2 99-171; efhand 94-122; efhand 146-174;
DEX0452_050.aa.2	Y	1 - i1- 4;tm5 - 27;o2 8- 363;	312- 318, 1.084; 5-28, 1.211; 249- 255, 1.102; 156- 161, 1.054; 132- 145, 1.09; 163- 169, 1.022; 34-43, 1.125; 274- 280, 1.072; 333- 338, 1.046; 50-68, 1.14; 114- 126, 1.063; 89- 102, 1.159; 343- 352, 1.062;	CK2_PHOSPHO_SITE 314- 317; PKC_PHOSPHO_SITE 322-324; MYRISTYL 238- 243; CK2_PHOSPHO_SITE 153-156; PKC_PHOSPHO_SITE 307- 309; CK2_PHOSPHO_SITE 307-310; PKC_PHOSPHO_SITE 85-87; MYRISTYL 347-352; CK2_PHOSPHO_SITE 172- 175; CK2_PHOSPHO_SITE 20-23; PKC_PHOSPHO_SITE 42-44; CK2_PHOSPHO_SITE 101-104; PKC_PHOSPHO_SITE 67-69; CK2_PHOSPHO_SITE 194- 197; TYR_PHOSPHO_SITE 272-278; CK2_PHOSPHO_SITE 205- 208; CK2_PHOSPHO_SITE 164-167; TYR_PHOSPHO_SITE 184- 191; TYR_PHOSPHO_SITE	EF_HAND 254- 266; efhand 245-273; EFh 245-273; EFh 297-325; efhand 297-325; EF_HAND_2 250- 322; EF_HAND 306-318;

				173-180; CK2_PHOSPHO_SITE 124- 127;	
DEX0452_051.aa.1	Y		89- 102, 1.159; 384- 395, 1.131; 312- 318, 1.084; 163- 169, 1.022; 1 - 132- ii- 145, 1.09; 4;tm5 156- - 161, 1.054; 27;o2 274- 8- 280, 1.072; 420; 114- 126, 1.063; 34-43, 1.125; 5-28, 1.211; 400- 409, 1.062; 50-68, 1.14; 249- 255, 1.102;	CK2_PHOSPHO_SITE 164- 167; CK2_PHOSPHO_SITE 172-175; CK2_PHOSPHO_SITE 205- 208; CK2_PHOSPHO_SITE 124-127; CK2_PHOSPHO_SITE 101- 104; CK2_PHOSPHO_SITE 194-197; TYR_PHOSPHO_SITE 173- 180; PKC_PHOSPHO_SITE 42-44; PKC_PHOSPHO_SITE 67-69; PKC_PHOSPHO_SITE 322-324; TYR_PHOSPHO_SITE 184- 191; PKC_PHOSPHO_SITE 307-309; TYR_PHOSPHO_SITE 272- 278; PKC_PHOSPHO_SITE 85-87; CK2_PHOSPHO_SITE 343-346; CK2_PHOSPHO_SITE 20-23; MYRISTYL 404-409; CK2_PHOSPHO_SITE 153- 156; CK2_PHOSPHO_SITE 307-310; MYRISTYL 238- 243; LEUCINE_ZIPPER 361-382; CK2_PHOSPHO_SITE 314- 317;	EFh 245-273; efhand 245-273; EFh 297-325; EF_HAND 254- 266; efhand 297-325; GLU_RICH 175- 374; EF_HAND 306-318; EF_HAND_2 250- 322;
DEX0452_052.aa.1	N		92-98, 1.088; 132- 138, 1.104; 12-21, 1.188; 26-40, 1.13; 0 - 117- o1- 130, 1.08; 162; 63-83, 1.101; 143- 158, 1.175; 100- 115, 1.089;	PKC_PHOSPHO_SITE 45-47; MYRISTYL 59-64; ASN GLYCOSYLATION 92- 95; PKC_PHOSPHO_SITE 48-50;	RASGAP_CTERM 24-106; sp_Q13576_IQG2_- HUMAN 24-159;
DEX0452_053.orf.1	N	0 - o1- 69;	54-66, 1.128; 4-16, 1.118; 21-48, 1.184;	PKC_PHOSPHO_SITE 51-53;	
DEX0452_053.aa.1	N	0 - o1- 42;	22-28, 1.06;	CK2_PHOSPHO_SITE 29-32; CK2_PHOSPHO_SITE 2-5;	
DEX0452_054.orf.1	N	1 - o1- 89;tm 88- 90- 112;i 113- 172;	159- 166, 1.149; 127, 1.252; 5-81, 1.262; 144- 150, 1.025;	CAMP_PHOSPHO_SITE 84- 87; MYRISTYL 61-66; CAMP_PHOSPHO_SITE 168- 171; ASN GLYCOSYLATION 86-89; CK2_PHOSPHO_SITE 154-157; MYRISTYL 17- 22; AMIDATION 81-84;	

DEX0452_054.aa.1	N	1 - ii- 50;tm 51- 73;o7 4- 134;	106- 112,1.025; 50-89,1.252; 5-20,1.069; 121- 128,1.149; 26-43,1.241;	MYRISTYL 28-33; CK2_PHOSPHO_SITE 116- 119; CAMP_PHOSPHO_SITE 130-133; ASN_GLYCOSYLATION 24- 27; MYRISTYL 23-28; ASN_GLYCOSYLATION 48- 51; CAMP_PHOSPHO_SITE 46-49; PKC_PHOSPHO_SITE 8-10; AMIDATION 43-46; ASN_GLYCOSYLATION 37- 40;	
DEX0452_055.aa.1	N	0 - o1- 396;	100- 106,1.096; 41-51,1.108; 182- 188,1.072; 89-97,1.175; 276- 333,1.174; 192- 197,1.061; 343- 372,1.126; 141- 161,1.131; 168- 180,1.158; 115- 138,1.154; 206- 214,1.115; 17-22,1.027; 61-70,1.094; 216- 272,1.139; 32-39,1.169;	MYRISTYL 66-71; PKC_PHOSPHO_SITE 133- 135; ASN_GLYCOSYLATION 13-16; PKC_PHOSPHO_SITE 27-29; TYR_PHOSPHO_SITE 201-208; MYRISTYL 307- 312; CK2_PHOSPHO_SITE 30-33; CAMP_PHOSPHO_SITE 376- 379; CK2_PHOSPHO_SITE 57-60; ASN_GLYCOSYLATION 28- 31; MYRISTYL 24-29; CK2_PHOSPHO_SITE 199- 202; PKC_PHOSPHO_SITE 199-201; PKC_PHOSPHO_SITE 15-17; CK2_PHOSPHO_SITE 77-80; PKC_PHOSPHO_SITE 47-49; MYRISTYL 218-223; PKC_PHOSPHO_SITE 70-72; MYRISTYL 152-157;	PSTLEXTEINS 320-343; PRICHEXTENSN 254-266; PRO_RICH 249- 371; PRICHEXTENSN 288-300; PSTLEXTEINS 353-371; PRICHEXTENSN 353-370; PRICHEXTENSN 313-334;
DEX0452_055.aa.2	N	0 - o1- 255;	65-73,1.115; 27-39,1.158; 202- 231,1.126; 51-56,1.061; 75- 131,1.139; 4-20,1.131; 135- 192,1.174; 41-47,1.072;	TYR_PHOSPHO_SITE 60-67; MYRISTYL 77-82; PKC_PHOSPHO_SITE 58-60; MYRISTYL 11-16; CAMP_PHOSPHO_SITE 235- 238; MYRISTYL 166-171; CK2_PHOSPHO_SITE 58-61;	PRICHEXTENSN 172-188; PSTLEXTEINS 212-230; PRICHEXTENSN 138-159; PRO_RICH 108- 230; PRICHEXTENSN 122-134; PRICHEXTENSN 204-229; PSTLEXTEINS 179-202;
DEX0452_056.orf.1	N	0 - o1- 412;	4-28,1.175; 148- 174,1.208; 340-347,1.1; 116- 121,1.101; 271- 278,1.142;	PKC_PHOSPHO_SITE 337- 339; PKC_PHOSPHO_SITE 83-85; PKC_PHOSPHO_SITE 191-193; CK2_PHOSPHO_SITE 73-76; CK2_PHOSPHO_SITE 284- 287; PKC_PHOSPHO_SITE 50-52; CK2_PHOSPHO SITE	PABP 26-62; PTS_HPR_SER 66- 81; HECT 99- 412; HECT 117- 412; HECTc 45- 412; PolyA 11- 65;

			60-65, 1.067; 319- 324, 1.064; 31-48, 1.19; 297- 307, 1.112; 349- 354, 1.068; 227- 251, 1.106; 254- 266, 1.144; 97- 103, 1.084; 192- 201, 1.122; 181- 190, 1.093; 288- 295, 1.117; 204- 222, 1.128; 73-86, 1.146; 372- 405, 1.183;	308-311; PKC_PHOSPHO_SITE 391- 393; CK2_PHOSPHO_SITE 329-332; AMIDATION 121- 124; PKC_PHOSPHO_SITE 366-368; CK2_PHOSPHO_SITE 335- 338; CK2_PHOSPHO_SITE 99-102; MYRISTYL 95- 100; PKC_PHOSPHO_SITE 128-130; CK2_PHOSPHO_SITE 279- 282; PKC_PHOSPHO_SITE 96-98; CK2_PHOSPHO_SITE 202-205; MYRISTYL 78- 83; AMIDATION 172-175; CK2_PHOSPHO_SITE 209- 212;	
DEX0452_056.aa.1	N	0 - o1- 56;	8-32, 1.169; 45-53, 1.094; 34-43, 1.131;	MYRISTYL 43-48; PKC_PHOSPHO_SITE 51-53; PKC_PHOSPHO_SITE 15-17; MYRISTYL 38-43;	
DEX0452_057.orf.1	N	0 - o1- 1 430;	129- 138, 1.145; 153- 170, 1.175; 392- 399, 1.102; 191- 215, 1.178; 17-27, 1.077; 178- 184, 1.041; 279- 289, 1.079; 140- 147, 1.14; 80-88, 1.094; 50-73, 1.227; 119- 125, 1.085; 335- 342, 1.111; 294- 328, 1.138; 30-36, 1.035; 220- 225, 1.042; 369- 386, 1.13; 96- 103.1.131:	MYRISTYL 29-34; AMIDATION 3-6; CAMP_PHOSPHO_SITE 275- 278; PKC_PHOSPHO_SITE 140-142; CK2_PHOSPHO_SITE 188- 191; CK2_PHOSPHO_SITE 419-422; ASN_GLYCOSYLATION 127- 130; AMIDATION 29-32; CK2_PHOSPHO_SITE 299- 302; PKC_PHOSPHO_SITE 80-82; CK2_PHOSPHO_SITE 135-138; CK2_PHOSPHO_SITE 89-92; AMIDATION 257-260; PKC_PHOSPHO_SITE 221- 223; PKC_PHOSPHO_SITE 121-123; MYRISTYL 17- 22; PKC_PHOSPHO_SITE 150-152; MYRISTYL 77- 82; MYRISTYL 25-30; PKC_PHOSPHO_SITE 279- 281; PKC_PHOSPHO_SITE 89-91; MYRISTYL 76-81; PKC_PHOSPHO_SITE 84-86; MYRISTYL 390-395; MYRISTYL 354-359; CK2_PHOSPHO_SITE 237- 240;	GPROTEINBRPT 351-365; GPROTEINBRPT 132-146; WD_REPEATS_REGI ON 194-397; WD40 188-226; WD40 230-265; GPROTEINBRPT 213-227; WD_REPEATS_2_2 339-364; WD40 104-145; WD40 316-364; WD40 148-185; WD40 190-226; WD40 326-364; WD40 229-265; sp_Q98UH2_Q98UH 2_XENLA 197- 226; WD40 367- 404; WD40 148- 185; WD40 106- 145; WD_REPEATS_2_1 194-226;

			239- 256, 1.159; 259- 265, 1.128;		
DEX0452_057.aa.1	N	0 - o1- 385;	5-28, 1.227; 74-80, 1.085; 214- 220, 1.128; 290- 297, 1.111; 194- 211, 1.159; 133- 139, 1.041; 146- 170, 1.178; 51-58, 1.131; 108- 125, 1.175; 249- 283, 1.138; 175- 180, 1.042; 324- 341, 1.13; 84-93, 1.145; 234- 244, 1.079; 347- 354, 1.102; 35-43, 1.094; 95-102, 1.14;	AMIDATION 212-215; PKC_PHOSPHO_SITE 105- 107; MYRISTYL 32-37; MYRISTYL 309-314; CK2_PHOSPHO_SITE 254- 257; CK2_PHOSPHO_SITE 44-47; PKC_PHOSPHO_SITE 44-46; PKC_PHOSPHO_SITE 234-236; PKC_PHOSPHO_SITE 95-97; CK2_PHOSPHO_SITE 192- 195; ASN GLYCOSYLATION 82-85; CK2_PHOSPHO_SITE 374-377; CK2_PHOSPHO_SITE 143- 146; PKC_PHOSPHO_SITE 76-78; PKC_PHOSPHO_SITE 35-37; PKC_PHOSPHO_SITE 176-178; MYRISTYL 345- 350; PKC_PHOSPHO_SITE 39-41; MYRISTYL 31-36; CAMP_PHOSPHO_SITE 230- 233; CK2_PHOSPHO_SITE 90-93;	WD_REPEATS_REGI ON 149-352; sp_Q98UH2_Q98UH 2_XENLA 152- 181; GPROTEINBRPT 87-101; WD40 185-220; WD40 184-220; WD40 143-181; GPROTEINBRPT 306-320; WD40 103-140; WD40 281-319; GPROTEINBRPT 168-182; WD40 103-140; WD40 59-100; WD40 61-100; WD_REPEATS_2_1 149-181; WD40 271-319; WD40 322-359; WD40 145-181; WD_REPEATS_2_2 294-319;
DEX0452_058.aa.1	Y	3 - i1- 6;tm7 - 28;o2 9- 55;tm 56- 78;i7 9- 90;tm 91- 113;o 114- 123;	89- 110, 1.254; 43-56, 1.154; 63-68, 1.11; 74-79, 1.092; 23-34, 1.178; 4-21, 1.287;	MYRISTYL 61-66; CK2_PHOSPHO_SITE 35-38;	
DEX0452_058.orf.2	N	0 - o1- 211;	41-67, 1.143; 105- 115, 1.231; 198- 208, 1.143; 4-13, 1.213; 121- 131, 1.054; 87-93, 1.094;	PKC_PHOSPHO_SITE 77-79; PKC_PHOSPHO_SITE 188- 190; AMIDATION 80-83; RGD 35-37; CK2_PHOSPHO_SITE 77-80; AMIDATION 99-102; AMIDATION 151-154; MYRISTYL 180-185;	
DEX0452_058.aa.2	N	0 - o1-	58-71, 1.043; 73-90, 1.171;	PKC_PHOSPHO_SITE 46-48; MYRISTYL 113-118;	ARG_RICH 32- 165: ATHOOK 33-

		178;	11-26,1.068; 104- 109,1.015; 95- 101,1.006; 164- 174,1.033; 28-34,1.043;	PKC_PHOSPHO_SITE 150- 152; MYRISTYL 114-119; PKC_PHOSPHO_SITE 121- 123; MYRISTYL 75-80; PKC_PHOSPHO_SITE 160- 162; PKC_PHOSPHO_SITE 49-51; MYRISTYL 166- 171; MYRISTYL 60-65; CAMP_PHOSPHO_SITE 102- 105; CAMP_PHOSPHO_SITE 147-150; MYRISTYL 12- 17;	43; ATHOOK 136- 146; ATHOOK 69- 80;
DEX0452_058.orf. 3	N	0 - o1- 211;	105- 115,1.231; 87-93,1.094; 41-67,1.143; 4-13,1.213; 198- 208,1.143; 121- 131,1.054;	CK2_PHOSPHO_SITE 77-80; AMIDATION 99-102; RGD 35-37; PKC_PHOSPHO_SITE 77-79; PKC_PHOSPHO_SITE 188-190; AMIDATION 151- 154; MYRISTYL 180-185; AMIDATION 80-83;	
DEX0452_058.orf. 4	N	0 - o1- 211;	4-13,1.213; 121- 131,1.054; 105- 115,1.231; 87-93,1.094; 41-67,1.143; 198- 208,1.143;	CK2_PHOSPHO_SITE 77-80; RGD 35-37; PKC_PHOSPHO_SITE 188- 190; AMIDATION 80-83; AMIDATION 151-154; MYRISTYL 180-185; PKC_PHOSPHO_SITE 77-79; AMIDATION 99-102;	
DEX0452_058.orf. 5	N	0 - o1- 211;	4-13,1.213; 105- 115,1.231; 121- 131,1.054; 87-93,1.094; 41-67,1.143; 198- 208,1.143;	AMIDATION 99-102; CK2_PHOSPHO_SITE 77-80; PKC_PHOSPHO_SITE 188- 190; RGD 35-37; AMIDATION 151-154; MYRISTYL 180-185; AMIDATION 80-83; PKC_PHOSPHO_SITE 77-79;	
DEX0452_058.orf. 6	N	0 - o1- 211;	41-67,1.143; 105- 115,1.231; 87-93,1.094; 121- 131,1.054; 4-13,1.213; 198- 208,1.143;	AMIDATION 80-83; PKC_PHOSPHO_SITE 188- 190; PKC_PHOSPHO_SITE 77-79; RGD 35-37; MYRISTYL 180-185; AMIDATION 151-154; CK2_PHOSPHO_SITE 77-80; AMIDATION 99-102;	
DEX0452_058.orf. 7	N	0 - o1- 211;	41-67,1.143; 198- 208,1.143; 4-13,1.213; 87-93,1.094; 105- 115,1.231; 121- 131,1.054;	PKC_PHOSPHO_SITE 77-79; CK2_PHOSPHO_SITE 77-80; PKC_PHOSPHO_SITE 188- 190; MYRISTYL 180-185; AMIDATION 99-102; AMIDATION 151-154; AMIDATION 80-83; RGD 35-37;	
DEX0452	N	0 -	105-	AMIDATION 151-154:	

058.orf. 8		01- 211;	115,1.231; 121- 131,1.054; 4-13,1.213; 87-93,1.094; 198- 208,1.143; 41-67,1.143;	CK2_PHOSPHO_SITE 77-80; AMIDATION 99-102; PKC_PHOSPHO_SITE 188- 190; RGD 35-37; AMIDATION 80-83; MYRISTYL 180-185; PKC_PHOSPHO_SITE 77-79;	
DEX0452- 058.orf. 9	N	0 - 01- 211;	105- 115,1.231; 121- 131,1.054; 87-93,1.094; 198- 208,1.143; 41-67,1.143; 4-13,1.213;	AMIDATION 99-102; RGD 35-37; AMIDATION 80-83; MYRISTYL 180-185; PKC_PHOSPHO_SITE 77-79; CK2_PHOSPHO_SITE 77-80; AMIDATION 151-154; PKC_PHOSPHO_SITE 188- 190;	

Example 1b: Sequence Alignment Support

Alignments between previously identified sequences and splice variant sequences are performed to confirm unique portions of splice variant nucleic acid and amino acid

5 sequences. The alignments are done using the Needle program in the European Molecular Biology Open Software Suite (EMBOSS) version 2.2.0 available at www.emboss.org from EMBnet (<http://www.emblnet.org>). Default settings are used unless otherwise noted. The Needle program in EMBOSS implements the Needleman-Wunsch algorithm.

Needleman, S. B., Wunsch, C. D., *J. Mol. Biol.* 48:443-453 (1970).

10 It is well known to those skilled in the art that implication of alignment algorithms by various programs may result in minor changes in the generated output. These changes include but are not limited to: alignment scores (percent identity, similarity, and gap), display of nonaligned flanking sequence regions, and number assignment to residues. These minor changes in the output of an alignment do not alter the physical characteristics
15 of the sequences or the differences between the sequences, e.g. regions of homology, insertions, or deletions.

Example 1c: RT-PCR Analysis

To detect the presence and tissue distribution of a particular splice variant Reverse Transcription-Polymerase Chain Reaction (RT-PCR) is performed using cDNA generated
20 from a panel of tissue RNAs. See, e.g., Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press (1989) and; Kawasaki ES *et al.*, *PNAS* 85(15):5698 (1988). Total RNA is extracted from a variety of tissues and first

strand cDNA is prepared with reverse transcriptase (RT). Each panel includes 23 cDNAs from five cancer types (lung, ovary, breast, colon, and prostate) and normal samples of testis, placenta and fetal brain. Each cancer set is composed of three cancer cDNAs from different donors and one normal pooled sample. Using a standard enzyme kit from BD Bioscience Clontech (Mountain View, CA), the target transcript is detected with sequence-specific primers designed to only amplify the particular splice variant. The PCR reaction is run on the GeneAmp PCR system 9700 (Applied Biosystem, Foster City, CA) thermocycler under optimal conditions. One of ordinary skill can design appropriate primers and determine optimal conditions. The amplified product is resolved on an agarose gel to detect a band of equivalent size to the predicted RT-PCR product. A band indicated the presence of the splice variant in a sample. The relation of the amplified product to the splice variant was subsequently confirmed by DNA sequencing.

After subcloning, all positively screened clones are sequence verified. The DNA sequence verification results show the splice variant contains the predicted sequence differences in comparison with the reference sequence.

Results for RT-PCR analysis in the table below include the sequence DEX ID, Lead Name, Cancer Tissue(s) the transcript was detected in, Normal Tissue(s) the transcript was detected in, the predicted length of the RT-PCR product, and the Confirmed Length of the RT-PCR product.

DEX ID	Lead Name	Cancer Tissue(s)	Normal Tissue(s)	Predicted Length	Confirmed Length
DEX0452_010.nt.1	Mam113	Lung, Ovary	None	747bp	747bp
DEX0452_033.nt.2	Mam128V3	Lung, Ovary, Breast, Colon, Prostate	Lung, Ovary, Breast, Colon, Prostate	286bp	286bp

20

RT-PCR results confirm the presence SEQ ID NO: 1-95 in biologic samples and distinguish between related transcripts.

Example 1d: Secretion Assay

To determine if a protein encoded by a splice variant is secreted from cells a secretion assay is preformed. A pcDNA3.1 clone containing the gene transcript which encodes the variant protein is transfected into 293T cells using the Superfect transfection reagent (Qiagen, Valencia CA). Transfected cells are incubated for 28 hours before the

media is collected and immediately spun down to remove any detached cells. The adherent cells are solubilized with lysis buffer (1% NP40, 10mM sodium phosphate pH7.0, and 0.15M NaCl). The lysed cells are collected and spun down and the supernatant extracted as cell lysate. Western immunoblot is carried out in the following manner: 15 μ l of the cell lysate and media are run on 4-12% NuPage Bis-Tris gel (Invitrogen, Carlsbad CA), and blotted onto a PVDF membrane (Invitrogen, Carlsbad CA). The blot is incubated with a polyclonal primary antibody which binds to the variant protein (Imgenex, San Diego CA) and polyclonal goat anti-rabbit-peroxidase secondary antibody (Sigma-Aldrich, St. Louis MO). The blot is developed with the ECL Plus chemiluminescent detection reagent (Amersham BioSciences, Piscataway NJ).

10 Secretion assay results are indicative of SEQ ID NO: 96-232 being a diagnostic marker and/or therapeutic target for cancer.

Example 2a: Gene Expression Analysis

Custom Microarray Experiment - Cancer

15 Custom oligonucleotide microarrays were provided by Agilent Technologies, Inc. (Palo Alto, CA). The microarrays were fabricated by Agilent using their technology for the *in-situ* synthesis of 60mer oligonucleotides (Hughes, et al. 2001, Nature Biotechnology 19:342-347). The 60mer microarray probes were designed by Agilent, from gene sequences provided by diaDexus, using Agilent proprietary algorithms. Whenever 20 possible two different 60mers were designed for each gene of interest.

All microarray experiments were two-color experiments and were preformed using Agilent-recommended protocols and reagents. Briefly, each microarray was hybridized with cRNAs synthesized from RNA (total RNA for ovarian and prostate, polyA+ RNA for lung, breast and colon samples), isolated from cancer and normal tissues, labeled with 25 fluorescent dyes Cyanine3 (Cy3) or Cyanine5 (Cy5) (NEN Life Science Products, Inc., Boston, MA) using a linear amplification method (Agilent). In each experiment the experimental sample was RNA isolated from cancer tissue from a single individual and the reference sample was a pool of RNA isolated from normal tissues of the same organ as the cancerous tissue (*i.e.* normal ovarian tissue in experiments with ovarian cancer 30 samples). Hybridizations were carried out at 60°C, overnight using Agilent *in-situ* hybridization buffer. Following washing, arrays were scanned with a GenePix 4000B

Microarray Scanner (Axon Instruments, Inc., Union City, CA). The resulting images were analyzed with GenePix Pro 3.0 Microarray Acquisition and Analysis Software (Axon).

Data normalization and expression profiling were done with Expressionist software from GeneData Inc. (Daly City, CA/Basel, Switzerland). Gene expression analysis was performed using only experiments that met certain quality criteria. The quality criteria that experiments must meet are a combination of evaluations performed by the Expressionist software and evaluations performed manually using raw and normalized data. To evaluate raw data quality, detection limits (the mean signal for a replicated negative control + 2 Standard Deviations (SD)) for each channel were calculated. The 10 detection limit is a measure of non-specific hybridization. Acceptable detection limits were defined for each dye (<80 for Cy5 and <150 for Cy3). Arrays with poor detection limits in one or both channels were not analyzed and the experiments were repeated. To evaluate normalized data quality, positive control elements included in the array were utilized. These array features should have a mean ratio of 1 (no differential expression). 15 If these features have a mean ratio of greater than 1.5-fold up or down, the experiments were not analyzed further and were repeated. In addition to traditional scatter plots demonstrating the distribution of signal in each experiment, the Expressionist software also has minimum thresholding criteria that employ user defined parameters to identify quality data. These thresholds include two distinct quality measurements: 1) minimum 20 area percentage, which is a measure of the integrity of each spot and 2) signal to noise ratio, which ensures that the signal being measured is significantly above any background (nonspecific) signal present. Only those features that met the threshold criteria were included in the filtering and analyses carried out by Expressionist. The thresholding settings employed require a minimum area percentage of 60% [(% pixels > background + 25 2SD)-(% pixels saturated)], and a minimum signal to noise ratio of 2.0 in both channels. By these criteria, very low expressors, saturated features and spots with abnormally high local background were not included in analysis.

Relative expression data was collected from Expressionist based on filtering and clustering analyses. Up-regulated genes were identified using criteria for the percentage 30 of experiments in which the gene is up-regulated by at least 2-fold. In general, up-regulation in ~30% of samples tested was used as a cutoff for filtering.

Two microarray experiments were preformed for each normal and cancer tissue pair. The tissue specific Array Chip for each cancer tissue is a unique microarray specific

to that tissue and cancer. The Multi-Cancer Array Chip is a universal microarray that was hybridized with samples from each of the cancers (ovarian, breast, colon, lung, and prostate). See the description below for the experiments specific to the different cancers.

Microarray Experiments and Data Tables

5 BREAST CANCER CHIPS

For breast cancer two different chip designs were evaluated with overlapping sets of a total of 36 samples, comparing the expression patterns of breast cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 10 normal breast tissues. For the Breast Array Chip, all 36 samples (9 stage I cancers, 23 stage II cancers, 4 stage III 10 cancers) were analyzed. These samples also represented 10 Grade 1/2 and 26 Grade 3 cancers. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, well differentiated; G2, moderately differentiated; G3, poorly differentiated; and G4, undifferentiated. AJCC Cancer Staging Handbook, pp. 9, (5th Ed, 1998). Samples were further grouped based on the expression patterns of the known breast cancer 15 associated genes Her2 and ER α (10 HER2 up, 26 HER2 not up, 20 ER up and 16 ER not up) and for the Multi-Cancer Array Chip, a subset of 20 of these samples (9 stage I cancers, 8 stage II cancers, 3 stage III cancers) were assessed.

The results for the statistically significant up-regulated genes on the Breast Array Chip are shown in Tables 1 and 2. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 3. The first two columns of each table contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all ("ALL") breast cancer samples, cancers corresponding to stage I ("ST1"), stages II and III ("ST2,3"), grades 1 and 2 ("GR1,2"), grade 3 ("GR3"), cancers exhibiting up-regulation of Her2 ("HER2up") 20 or ER α ("ERup") or those not exhibiting up-regulation of Her2 ("NOT HER2up") or ER α ("NOT ERup"). "%up" indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=36 for Colon Array Chip, n=20 for the Multi-Cancer Array Chip), "%valid up" indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

25 30 Table 1.

DEX ID	Oligo Name	Mam ALL % up n=36	Mam ALL % valid up n=36	Mam ST1 % up n=9	Mam ST1 % valid up n=9	Mam ST2, 3 % up n=27	Mam ST2, 3 % valid up n=27	Mam GR1, 2 % up n=10	Mam GR1, 2 % valid up n=10	Mam GR3 % up n=26	Mam GR3 % valid up n=26
DEX0452_001.nt.1	34132.0	33.3	35.3	44.4	44.4	29.6	32.0	80.0	80.0	15.4	16.7
DEX0452_001.nt.1	34133.0	30.6	35.5	44.4	57.1	25.9	29.2	80.0	80.0	11.5	14.3
DEX0452_002.nt.1	13283.0	11.1	28.6	11.1	33.3	11.1	27.3	30.0	33.3	3.8	20.0
DEX0452_002.nt.1	13284.0	11.1	21.1	11.1	33.3	11.1	18.8	30.0	37.5	3.8	9.1
DEX0452_003.nt.1	14380.0	44.4	44.4	55.6	55.6	40.7	40.7	40.0	40.0	46.2	46.2
DEX0452_003.nt.1	14381.0	38.9	42.4	55.6	55.6	33.3	37.5	40.0	44.4	38.5	41.7
DEX0452_003.nt.2	14380.0	44.4	44.4	55.6	55.6	40.7	40.7	40.0	40.0	46.2	46.2
DEX0452_003.nt.2	14381.0	38.9	42.4	55.6	55.6	33.3	37.5	40.0	44.4	38.5	41.7
DEX0452_004.nt.1	28910.0	8.3	8.3	11.1	11.1	7.4	7.4	30.0	30.0	0.0	0.0
DEX0452_005.nt.1	16289.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_005.nt.1	16290.0	2.8	4.3	0.0	0.0	3.7	5.9	0.0	0.0	3.8	7.1
DEX0452_005.nt.1	29727.0	16.7	27.3	44.4	66.7	7.4	12.5	30.0	33.3	11.5	23.1
DEX0452_005.nt.1	29728.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_006.nt.1	20369.0	5.6	5.6	22.2	22.2	0.0	0.0	10.0	10.0	3.8	3.8
DEX0452_006.nt.1	20370.0	2.8	2.9	11.1	11.1	0.0	0.0	0.0	0.0	3.8	4.0
DEX0452_007.nt.1	12615.0	13.9	13.9	33.3	33.3	7.4	7.4	20.0	20.0	11.5	11.5
DEX0452_007.nt.1	12616.0	8.3	8.6	22.2	22.2	3.7	3.8	10.0	10.0	7.7	8.0
DEX0452_008.nt.1	27530.0	30.6	30.6	22.2	22.2	33.3	33.3	40.0	40.0	26.9	26.9
DEX0452_009.nt.1	20207.0	19.4	20.0	11.1	12.5	22.2	22.2	20.0	20.0	19.2	20.0
DEX0452_009.nt.2	20208.0	25.0	25.0	11.1	11.1	29.6	29.6	30.0	30.0	23.1	23.1
DEX0452_010.nt.1	15032.0	27.8	27.8	33.3	33.3	25.9	25.9	20.0	20.0	30.8	30.8
DEX0452_010.nt.1	15033.0	33.3	33.3	44.4	44.4	29.6	29.6	40.0	40.0	30.8	30.8
DEX0452_010.nt.1	31614.0	36.1	37.1	44.4	44.4	33.3	34.6	30.0	30.0	38.5	40.0
DEX0452_010.nt.1	31615.0	30.6	30.6	44.4	44.4	25.9	25.9	30.0	30.0	30.8	30.8
DEX0452_011.nt.1	31927.0	22.2	22.2	22.2	22.2	22.2	22.2	20.0	20.0	23.1	23.1
DEX0452_013.nt.1	11156.0	25.0	26.5	22.2	25.0	25.9	26.9	80.0	80.0	3.8	4.2

DEX0452_014.nt.1	38921.0	19.4	23.3	0.0	0.0	25.9	30.4	10.0	10.0	23.1	30.0
DEX0452_014.nt.1	38922.0	19.4	28.0	0.0	0.0	25.9	36.8	10.0	10.0	23.1	40.0
DEX0452_015.nt.1	18118.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.1	18250.0	30.6	30.6	44.4	44.4	25.9	25.9	20.0	20.0	34.6	34.6
DEX0452_015.nt.1	18256.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.2	18118.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.2	18250.0	30.6	30.6	44.4	44.4	25.9	25.9	20.0	20.0	34.6	34.6
DEX0452_015.nt.2	18256.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.3	18118.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.3	18250.0	30.6	30.6	44.4	44.4	25.9	25.9	20.0	20.0	34.6	34.6
DEX0452_015.nt.3	18256.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.4	18118.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.4	18250.0	30.6	30.6	44.4	44.4	25.9	25.9	20.0	20.0	34.6	34.6
DEX0452_015.nt.4	18256.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.5	18118.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_015.nt.5	18250.0	30.6	30.6	44.4	44.4	25.9	25.9	20.0	20.0	34.6	34.6
DEX0452_015.nt.5	18256.0	13.9	13.9	11.1	11.1	14.8	14.8	10.0	10.0	15.4	15.4
DEX0452_016.nt.1	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0
DEX0452_016.nt.1	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.1	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_016.nt.2	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0
DEX0452_016.nt.2	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.2	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.2	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.2	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_016.nt.3	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0
DEX0452_016.nt.3	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.3	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_016.nt.4	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0

DEX0452_016.nt.4	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.4	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_016.nt.5	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0
DEX0452_016.nt.5	19497.0	8.3	8.3	11.1	11.1	7.4	7.4	10.0	10.0	7.7	7.7
DEX0452_016.nt.5	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.5	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.5	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.5	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_016.nt.6	19496.0	11.1	13.3	11.1	12.5	11.1	13.6	10.0	10.0	11.5	15.0
DEX0452_016.nt.6	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.6	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.6	40273.0	11.1	13.8	11.1	11.1	11.1	15.0	10.0	10.0	11.5	15.8
DEX0452_016.nt.6	40284.0	5.6	5.9	11.1	11.1	3.7	4.0	0.0	0.0	7.7	8.0
DEX0452_017.nt.1	25674.0	22.2	25.8	11.1	14.3	25.9	29.2	30.0	30.0	19.2	23.8
DEX0452_017.nt.1	25675.0	19.4	22.6	11.1	11.1	22.2	27.3	20.0	20.0	19.2	23.8
DEX0452_018.nt.1	21561.0	22.2	32.0	22.2	40.0	22.2	30.0	50.0	50.0	11.5	20.0
DEX0452_018.nt.1	21562.0	22.2	28.6	22.2	33.3	22.2	27.3	50.0	50.0	11.5	16.7
DEX0452_019.nt.1	12953.0	22.2	22.2	22.2	22.2	22.2	22.2	40.0	40.0	15.4	15.4
DEX0452_019.nt.1	12954.0	22.2	22.2	22.2	22.2	22.2	22.2	40.0	40.0	15.4	15.4
DEX0452_020.nt.1	17932.0	36.1	37.1	44.4	44.4	33.3	34.6	50.0	50.0	30.8	32.0
DEX0452_020.nt.1	17933.0	38.9	38.9	33.3	33.3	40.7	40.7	50.0	50.0	34.6	34.6
DEX0452_020.nt.1	17934.0	30.6	31.4	33.3	33.3	29.6	30.8	40.0	40.0	26.9	28.0
DEX0452_020.nt.1	17938.0	36.1	38.2	33.3	33.3	37.0	40.0	50.0	50.0	30.8	33.3
DEX0452_020.nt.1	17942.0	36.1	36.1	33.3	33.3	37.0	37.0	50.0	50.0	30.8	30.8
DEX0452_021.nt.1	25824.0	30.6	33.3	44.4	44.4	25.9	29.2	60.0	60.0	19.2	21.7
DEX0452_022.nt.1	29793.0	50.0	72.0	55.6	83.3	48.1	68.4	90.0	90.0	34.6	60.0
DEX0452_022.nt.1	29794.0	47.2	68.0	44.4	80.0	48.1	65.0	90.0	90.0	30.8	53.3
DEX0452_023.nt.1	19174.0	11.1	12.9	11.1	14.3	11.1	12.5	30.0	30.0	3.8	4.8
DEX0452_023.nt.1	19175.0	5.6	5.6	0.0	0.0	7.4	7.4	10.0	10.0	3.8	3.8

DEX0452_024.nt.1	13892.0	22.2	22.2	11.1	11.1	25.9	25.9	0.0	0.0	30.8	30.8
DEX0452_025.nt.1	18383.0	25.0	27.3	33.3	37.5	22.2	24.0	40.0	40.0	19.2	21.7
DEX0452_026.nt.1	35953.0	27.8	41.7	55.6	71.4	18.5	29.4	70.0	77.8	11.5	20.0
DEX0452_026.nt.1	35954.0	19.4	21.9	33.3	37.5	14.8	16.7	50.0	55.6	7.7	8.7
DEX0452_027.nt.1	33040.0	19.4	19.4	0.0	0.0	25.9	25.9	20.0	20.0	19.2	19.2
DEX0452_027.nt.1	33041.0	8.3	8.3	0.0	0.0	11.1	11.1	10.0	10.0	7.7	7.7
DEX0452_027.nt.2	33040.0	19.4	19.4	0.0	0.0	25.9	25.9	20.0	20.0	19.2	19.2
DEX0452_027.nt.2	33041.0	8.3	8.3	0.0	0.0	11.1	11.1	10.0	10.0	7.7	7.7
DEX0452_029.nt.1	19254.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_029.nt.1	19255.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_029.nt.1	33276.0	27.8	29.4	33.3	37.5	25.9	26.9	10.0	10.0	34.6	37.5
DEX0452_029.nt.1	33277.0	16.7	17.1	22.2	22.2	14.8	15.4	10.0	10.0	19.2	20.0
DEX0452_029.nt.2	33276.0	27.8	29.4	33.3	37.5	25.9	26.9	10.0	10.0	34.6	37.5
DEX0452_029.nt.2	33277.0	16.7	17.1	22.2	22.2	14.8	15.4	10.0	10.0	19.2	20.0
DEX0452_030.nt.1	27825.0	13.9	14.3	0.0	0.0	18.5	18.5	30.0	30.0	7.7	8.0
DEX0452_030.nt.1	27826.0	13.9	13.9	0.0	0.0	18.5	18.5	30.0	30.0	7.7	7.7
DEX0452_031.nt.1	32496.0	8.3	9.1	33.3	33.3	0.0	0.0	20.0	20.0	3.8	4.3
DEX0452_031.nt.1	32497.0	8.3	8.3	33.3	33.3	0.0	0.0	20.0	20.0	3.8	3.8
DEX0452_031.nt.2	32496.0	8.3	9.1	33.3	33.3	0.0	0.0	20.0	20.0	3.8	4.3
DEX0452_031.nt.2	32497.0	8.3	8.3	33.3	33.3	0.0	0.0	20.0	20.0	3.8	3.8
DEX0452_031.nt.3	32496.0	8.3	9.1	33.3	33.3	0.0	0.0	20.0	20.0	3.8	4.3
DEX0452_031.nt.3	32497.0	8.3	8.3	33.3	33.3	0.0	0.0	20.0	20.0	3.8	3.8
DEX0452_032.nt.1	31576.0	2.8	5.3	11.1	20.0	0.0	0.0	0.0	0.0	3.8	9.1
DEX0452_032.nt.1	31577.0	5.6	6.5	22.2	28.6	0.0	0.0	10.0	10.0	3.8	4.8
DEX0452_032.nt.1	40320.0	11.1	11.1	33.3	33.3	3.7	3.7	20.0	20.0	7.7	7.7
DEX0452_032.nt.1	40363.0	11.1	11.1	33.3	33.3	3.7	3.7	20.0	20.0	7.7	7.7
DEX0452_032.nt.1	40364.0	11.1	11.1	33.3	33.3	3.7	3.7	20.0	20.0	7.7	7.7
DEX0452_034.nt.1	25930.0	36.1	36.1	11.1	11.1	44.4	44.4	30.0	30.0	38.5	38.5
DEX0452_034.nt.1	25931.0	33.3	33.3	11.1	11.1	40.7	40.7	30.0	30.0	34.6	34.6

DEX0452_034.nt.2	25930.0	36.1	36.1	11.1	11.1	44.4	44.4	30.0	30.0	38.5	38.5
DEX0452_034.nt.2	25931.0	33.3	33.3	11.1	11.1	40.7	40.7	30.0	30.0	34.6	34.6
DEX0452_034.nt.3	25930.0	36.1	36.1	11.1	11.1	44.4	44.4	30.0	30.0	38.5	38.5
DEX0452_034.nt.3	25931.0	33.3	33.3	11.1	11.1	40.7	40.7	30.0	30.0	34.6	34.6
DEX0452_035.nt.1	27220.0	16.7	16.7	11.1	11.1	18.5	18.5	30.0	30.0	11.5	11.5
DEX0452_036.nt.1	27219.0	11.1	11.1	11.1	11.1	11.1	11.1	20.0	20.0	7.7	7.7
DEX0452_036.nt.1	27220.0	16.7	16.7	11.1	11.1	18.5	18.5	30.0	30.0	11.5	11.5
DEX0452_036.nt.2	27219.0	11.1	11.1	11.1	11.1	11.1	11.1	20.0	20.0	7.7	7.7
DEX0452_036.nt.2	27220.0	16.7	16.7	11.1	11.1	18.5	18.5	30.0	30.0	11.5	11.5
DEX0452_037.nt.1	27233.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_037.nt.1	27234.0	16.7	16.7	22.2	22.2	14.8	14.8	0.0	0.0	23.1	23.1
DEX0452_037.nt.1	40267.0	2.8	2.9	11.1	12.5	0.0	0.0	10.0	10.0	0.0	0.0
DEX0452_037.nt.2	27233.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_037.nt.2	27234.0	16.7	16.7	22.2	22.2	14.8	14.8	0.0	0.0	23.1	23.1
DEX0452_038.nt.1	40103.0	33.3	33.3	0.0	0.0	44.4	44.4	40.0	40.0	30.8	30.8
DEX0452_039.nt.1	12621.0	13.9	14.3	11.1	11.1	14.8	15.4	10.0	10.0	15.4	16.0
DEX0452_039.nt.1	12622.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_039.nt.1	12631.0	55.6	58.8	55.6	55.6	55.6	60.0	50.0	55.6	57.7	60.0
DEX0452_039.nt.1	27217.0	61.1	62.9	55.6	55.6	63.0	65.4	60.0	60.0	61.5	64.0
DEX0452_039.nt.1	27218.0	61.1	62.9	55.6	55.6	63.0	65.4	60.0	60.0	61.5	64.0
DEX0452_040.nt.1	24442.0	25.0	25.0	11.1	11.1	29.6	29.6	10.0	10.0	30.8	30.8
DEX0452_040.nt.1	24443.0	16.7	26.1	11.1	16.7	18.5	29.4	0.0	0.0	23.1	33.3
DEX0452_041.nt.1	20612.0	25.0	25.0	11.1	11.1	29.6	29.6	40.0	40.0	19.2	19.2
DEX0452_042.nt.1	27229.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	27230.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_043.nt.1	28899.0	8.3	8.3	22.2	22.2	3.7	3.7	20.0	20.0	3.8	3.8
DEX0452_044.nt.1	27063.0	22.2	38.1	33.3	75.0	18.5	29.4	60.0	60.0	7.7	18.2
DEX0452_044.nt.1	27064.0	11.1	28.6	33.3	100.0	3.7	9.1	40.0	44.4	0.0	0.0
DEX0452_044.nt.2	27063.0	22.2	38.1	33.3	75.0	18.5	29.4	60.0	60.0	7.7	18.2

DEX0452_044.nt.2	27064.0	11.1	28.6	33.3	100.0	3.7	9.1	40.0	44.4	0.0	0.0
DEX0452_045.nt.1	30175.0	41.7	48.4	33.3	50.0	44.4	48.0	60.0	66.7	34.6	40.9
DEX0452_045.nt.1	30176.0	50.0	66.7	33.3	50.0	55.6	71.4	60.0	75.0	46.2	63.2
DEX0452_046.nt.1	20370.0	2.8	2.9	11.1	11.1	0.0	0.0	0.0	0.0	3.8	4.0
DEX0452_046.nt.2	20369.0	5.6	5.6	22.2	22.2	0.0	0.0	10.0	10.0	3.8	3.8
DEX0452_046.nt.2	20370.0	2.8	2.9	11.1	11.1	0.0	0.0	0.0	0.0	3.8	4.0
DEX0452_047.nt.1	34092.0	25.0	25.7	22.2	22.2	25.9	26.9	20.0	20.0	26.9	28.0
DEX0452_048.nt.1	26236.0	19.4	20.0	0.0	0.0	25.9	26.9	20.0	20.0	19.2	20.0
DEX0452_048.nt.1	26237.0	13.9	15.6	0.0	0.0	18.5	20.8	10.0	10.0	15.4	18.2
DEX0452_049.nt.1	40305.0	11.1	12.1	11.1	12.5	11.1	12.0	30.0	30.0	3.8	4.3
DEX0452_049.nt.1	40306.0	33.3	33.3	33.3	33.3	33.3	33.3	50.0	50.0	26.9	26.9
DEX0452_049.nt.2	40305.0	11.1	12.1	11.1	12.5	11.1	12.0	30.0	30.0	3.8	4.3
DEX0452_049.nt.2	40306.0	33.3	33.3	33.3	33.3	33.3	33.3	50.0	50.0	26.9	26.9
DEX0452_050.nt.1	19465.0	41.7	41.7	33.3	33.3	44.4	44.4	80.0	80.0	26.9	26.9
DEX0452_052.nt.1	29054.0	13.9	14.7	22.2	25.0	11.1	11.5	30.0	30.0	7.7	8.3
DEX0452_053.nt.1	41778.0	8.3	8.8	22.2	22.2	3.7	4.0	30.0	30.0	0.0	0.0
DEX0452_054.nt.1	27617.0	16.7	16.7	33.3	33.3	11.1	11.1	30.0	30.0	11.5	11.5
DEX0452_054.nt.1	27618.0	13.9	13.9	22.2	22.2	11.1	11.1	20.0	20.0	11.5	11.5
DEX0452_055.nt.1	22448.0	25.0	25.0	33.3	33.3	22.2	22.2	30.0	30.0	23.1	23.1
DEX0452_056.nt.1	14317.0	27.8	27.8	22.2	22.2	29.6	29.6	40.0	40.0	23.1	23.1
DEX0452_056.nt.1	15115.0	2.8	5.9	0.0	0.0	3.7	7.7	10.0	12.5	0.0	0.0
DEX0452_056.nt.1	26101.0	22.2	22.2	22.2	22.2	22.2	22.2	30.0	30.0	19.2	19.2
DEX0452_057.nt.1	24447.0	22.2	22.2	33.3	33.3	18.5	18.5	50.0	50.0	11.5	11.5
DEX0452_058.nt.2	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.2	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.3	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.3	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.4	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.4	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9

DEX0452_058.nt.5	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.5	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.6	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.6	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.7	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.7	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.8	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.8	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9
DEX0452_058.nt.9	30041.0	38.9	38.9	55.6	55.6	33.3	33.3	40.0	40.0	38.5	38.5
DEX0452_058.nt.9	30042.0	25.0	25.0	44.4	44.4	18.5	18.5	20.0	20.0	26.9	26.9

Table 2.

DEX ID	Oligo Name	Mam HER2up %up n=10	Mam HER2up %valid n=10	Mam NOT HER2up %up n=26	Mam NOT HER2up %valid n=26	Mam ERup %up n=20	Mam ERup %valid n=20	Mam NOT ERup %up n=16	Mam NOT ERup %valid n=16
DEX0452_001.nt.1	34132.0	60.0	60.0	23.1	25.0	50.0	50.0	12.5	14.3
DEX0452_001.nt.1	34133.0	50.0	55.6	23.1	27.3	45.0	47.4	12.5	16.7
DEX0452_002.nt.1	13283.0	10.0	20.0	11.5	33.3	20.0	36.4	0.0	0.0
DEX0452_002.nt.1	13284.0	10.0	20.0	11.5	21.4	20.0	30.8	0.0	0.0
DEX0452_003.nt.1	14380.0	30.0	30.0	50.0	50.0	30.0	30.0	62.5	62.5
DEX0452_003.nt.1	14381.0	30.0	33.3	42.3	45.8	30.0	33.3	50.0	53.3
DEX0452_003.nt.2	14380.0	30.0	30.0	50.0	50.0	30.0	30.0	62.5	62.5
DEX0452_003.nt.2	14381.0	30.0	33.3	42.3	45.8	30.0	33.3	50.0	53.3
DEX0452_004.nt.1	28910.0	10.0	10.0	7.7	7.7	15.0	15.0	0.0	0.0
DEX0452_005.nt.1	16289.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_005.nt.1	16290.0	0.0	0.0	3.8	5.9	5.0	6.2	0.0	0.0
DEX0452_005.nt.1	29727.0	0.0	0.0	23.1	37.5	30.0	33.3	0.0	0.0
DEX0452_005.nt.1	29728.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_006.nt.1	20369.0	0.0	0.0	7.7	7.7	5.0	5.0	6.2	6.2
DEX0452_006.nt.1	20370.0	0.0	0.0	3.8	4.0	0.0	0.0	6.2	6.2
DEX0452_007.nt.1	12615.0	10.0	10.0	15.4	15.4	15.0	15.0	12.5	12.5
DEX0452_007.nt.1	12616.0	0.0	0.0	11.5	12.0	10.0	10.5	6.2	6.2
DEX0452_008.nt.1	27530.0	40.0	40.0	26.9	26.9	20.0	20.0	43.8	43.8
DEX0452_009.nt.1	20207.0	70.0	70.0	0.0	0.0	15.0	15.8	25.0	25.0
DEX0452_009.nt.2	20208.0	90.0	90.0	0.0	0.0	25.0	25.0	25.0	25.0
DEX0452_010.nt.1	15032.0	30.0	30.0	26.9	26.9	10.0	10.0	50.0	50.0
DEX0452_010.nt.1	15033.0	30.0	30.0	34.6	34.6	20.0	20.0	50.0	50.0
DEX0452_010.nt.1	31614.0	40.0	44.4	34.6	34.6	20.0	21.1	56.2	56.2
DEX0452_010.nt.1	31615.0	40.0	40.0	26.9	26.9	15.0	15.0	50.0	50.0
DEX0452_011.nt.1	31927.0	20.0	20.0	23.1	23.1	20.0	20.0	25.0	25.0
DEX0452_013.nt.1	11156.0	40.0	44.4	19.2	20.0	40.0	42.1	6.2	6.7
DEX0452_014.nt.1	38921.0	30.0	33.3	15.4	19.0	20.0	22.2	18.8	25.0
DEX0452_014.nt.1	38922.0	30.0	33.3	15.4	25.0	20.0	23.5	18.8	37.5
DEX0452_015.nt.1	18118.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.1	18250.0	30.0	30.0	30.8	30.8	30.0	30.0	31.2	31.2

DEX0452_015.nt.1	18256.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.2	18118.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.2	18250.0	30.0	30.0	30.8	30.8	30.0	30.0	31.2	31.2
DEX0452_015.nt.2	18256.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.3	18118.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.3	18250.0	30.0	30.0	30.8	30.8	30.0	30.0	31.2	31.2
DEX0452_015.nt.3	18256.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.4	18118.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.4	18250.0	30.0	30.0	30.8	30.8	30.0	30.0	31.2	31.2
DEX0452_015.nt.4	18256.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.5	18118.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_015.nt.5	18250.0	30.0	30.0	30.8	30.8	30.0	30.0	31.2	31.2
DEX0452_015.nt.5	18256.0	20.0	20.0	11.5	11.5	15.0	15.0	12.5	12.5
DEX0452_016.nt.1	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.1	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.1	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_016.nt.2	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.2	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.2	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.2	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.2	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_016.nt.3	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.3	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.3	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_016.nt.4	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.4	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.4	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_016.nt.5	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.5	19497.0	20.0	20.0	3.8	3.8	5.0	5.0	12.5	12.5
DEX0452_016.nt.5	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.5	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.5	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.5	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_016.nt.6	19496.0	30.0	42.9	3.8	4.3	5.0	5.9	18.8	23.1
DEX0452_016.nt.6	20285.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.6	20286.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_016.nt.6	40273.0	30.0	42.9	3.8	4.5	5.0	5.6	18.8	27.3
DEX0452_016.nt.6	40284.0	20.0	22.2	0.0	0.0	0.0	0.0	12.5	13.3
DEX0452_017.nt.1	25674.0	10.0	12.5	26.9	30.4	25.0	29.4	18.8	21.4
DEX0452_017.nt.1	25675.0	10.0	14.3	23.1	25.0	20.0	22.2	18.8	23.1
DEX0452_018.nt.1	21561.0	10.0	14.3	26.9	38.9	35.0	35.0	6.2	20.0
DEX0452_018.nt.1	21562.0	10.0	14.3	26.9	33.3	35.0	35.0	6.2	12.5
DEX0452_019.nt.1	12953.0	10.0	10.0	26.9	26.9	25.0	25.0	18.8	18.8
DEX0452_019.nt.1	12954.0	10.0	10.0	26.9	26.9	25.0	25.0	18.8	18.8
DEX0452_020.nt.1	17932.0	30.0	30.0	38.5	40.0	40.0	40.0	31.2	33.3
DEX0452_020.nt.1	17933.0	40.0	40.0	38.5	38.5	40.0	40.0	37.5	37.5
DEX0452_020.nt.1	17934.0	20.0	22.2	34.6	34.6	30.0	30.0	31.2	33.3
DEX0452_020.nt.1	17938.0	30.0	33.3	38.5	40.0	35.0	38.9	37.5	37.5
DEX0452_020.nt.1	17942.0	30.0	30.0	38.5	38.5	35.0	35.0	37.5	37.5
DEX0452_021.nt.1	25824.0	0.0	0.0	42.3	45.8	45.0	45.0	12.5	15.4
DEX0452_022.nt.1	29793.0	70.0	77.8	42.3	68.8	80.0	80.0	12.5	40.0
DEX0452_022.nt.1	29794.0	80.0	88.9	34.6	56.2	75.0	83.3	12.5	28.6
DEX0452_023.nt.1	19174.0	10.0	12.5	11.5	13.0	15.0	17.6	6.2	7.1
DEX0452_023.nt.1	19175.0	10.0	10.0	3.8	3.8	5.0	5.0	6.2	6.2
DEX0452_024.nt.1	13892.0	20.0	20.0	23.1	23.1	5.0	5.0	43.8	43.8
DEX0452_025.nt.1	18383.0	40.0	50.0	19.2	20.0	30.0	30.0	18.8	23.1
DEX0452_026.nt.1	35953.0	20.0	33.3	30.8	44.4	35.0	50.0	18.8	30.0

DEX0452_026.nt.1	35954.0	20.0	25.0	19.2	20.8	25.0	26.3	12.5	15.4
DEX0452_027.nt.1	33040.0	20.0	20.0	19.2	19.2	25.0	25.0	12.5	12.5
DEX0452_027.nt.1	33041.0	10.0	10.0	7.7	7.7	10.0	10.0	6.2	6.2
DEX0452_027.nt.2	33040.0	20.0	20.0	19.2	19.2	25.0	25.0	12.5	12.5
DEX0452_027.nt.2	33041.0	10.0	10.0	7.7	7.7	10.0	10.0	6.2	6.2
DEX0452_029.nt.1	19254.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_029.nt.1	19255.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_029.nt.1	33276.0	30.0	30.0	26.9	29.2	20.0	21.1	37.5	40.0
DEX0452_029.nt.1	33277.0	10.0	11.1	19.2	19.2	15.0	15.8	18.8	18.8
DEX0452_029.nt.2	33276.0	30.0	30.0	26.9	29.2	20.0	21.1	37.5	40.0
DEX0452_029.nt.2	33277.0	10.0	11.1	19.2	19.2	15.0	15.8	18.8	18.8
DEX0452_030.nt.1	27825.0	40.0	40.0	3.8	4.0	20.0	20.0	6.2	6.7
DEX0452_030.nt.1	27826.0	40.0	40.0	3.8	3.8	20.0	20.0	6.2	6.2
DEX0452_031.nt.1	32496.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	7.7
DEX0452_031.nt.1	32497.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	6.2
DEX0452_031.nt.2	32496.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	7.7
DEX0452_031.nt.2	32497.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	6.2
DEX0452_031.nt.3	32496.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	7.7
DEX0452_031.nt.3	32497.0	0.0	0.0	11.5	11.5	10.0	10.0	6.2	6.2
DEX0452_032.nt.1	31576.0	0.0	0.0	3.8	6.7	0.0	0.0	6.2	12.5
DEX0452_032.nt.1	31577.0	0.0	0.0	7.7	8.7	5.0	5.6	6.2	7.7
DEX0452_032.nt.1	40320.0	0.0	0.0	15.4	15.4	15.0	15.0	6.2	6.2
DEX0452_032.nt.1	40363.0	0.0	0.0	15.4	15.4	15.0	15.0	6.2	6.2
DEX0452_032.nt.1	40364.0	0.0	0.0	15.4	15.4	15.0	15.0	6.2	6.2
DEX0452_034.nt.1	25930.0	100.0	100.0	11.5	11.5	45.0	45.0	25.0	25.0
DEX0452_034.nt.1	25931.0	100.0	100.0	7.7	7.7	40.0	40.0	25.0	25.0
DEX0452_034.nt.2	25930.0	100.0	100.0	11.5	11.5	45.0	45.0	25.0	25.0
DEX0452_034.nt.2	25931.0	100.0	100.0	7.7	7.7	40.0	40.0	25.0	25.0
DEX0452_034.nt.3	25930.0	100.0	100.0	11.5	11.5	45.0	45.0	25.0	25.0
DEX0452_034.nt.3	25931.0	100.0	100.0	7.7	7.7	40.0	40.0	25.0	25.0
DEX0452_035.nt.1	27220.0	10.0	10.0	19.2	19.2	25.0	25.0	6.2	6.2
DEX0452_036.nt.1	27219.0	0.0	0.0	15.4	15.4	15.0	15.0	6.2	6.2
DEX0452_036.nt.1	27220.0	10.0	10.0	19.2	19.2	25.0	25.0	6.2	6.2
DEX0452_036.nt.2	27219.0	0.0	0.0	15.4	15.4	15.0	15.0	6.2	6.2
DEX0452_036.nt.2	27220.0	10.0	10.0	19.2	19.2	25.0	25.0	6.2	6.2
DEX0452_037.nt.1	27233.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_037.nt.1	27234.0	20.0	20.0	15.4	15.4	5.0	5.0	31.2	31.2
DEX0452_037.nt.1	40267.0	0.0	0.0	3.8	4.0	5.0	5.0	0.0	0.0
DEX0452_037.nt.2	27233.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_037.nt.2	27234.0	20.0	20.0	15.4	15.4	5.0	5.0	31.2	31.2
DEX0452_038.nt.1	40103.0	40.0	40.0	30.8	30.8	35.0	35.0	31.2	31.2
DEX0452_039.nt.1	12621.0	10.0	10.0	15.4	16.0	15.0	15.8	12.5	12.5
DEX0452_039.nt.1	12622.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_039.nt.1	12631.0	50.0	62.5	57.7	57.7	50.0	52.6	62.5	66.7
DEX0452_039.nt.1	127217.0	50.0	55.6	65.4	65.4	55.0	55.0	68.8	73.3
DEX0452_039.nt.1	127218.0	60.0	66.7	61.5	61.5	55.0	55.0	68.8	73.3
DEX0452_040.nt.1	24442.0	20.0	20.0	26.9	26.9	5.0	5.0	50.0	50.0
DEX0452_040.nt.1	24443.0	0.0	0.0	23.1	33.3	0.0	0.0	37.5	60.0
DEX0452_041.nt.1	20612.0	20.0	20.0	26.9	26.9	25.0	25.0	25.0	25.0
DEX0452_042.nt.1	27229.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	27230.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_043.nt.1	28899.0	0.0	0.0	11.5	11.5	15.0	15.0	0.0	0.0
DEX0452_044.nt.1	27063.0	10.0	16.7	26.9	46.7	30.0	42.9	12.5	28.6
DEX0452_044.nt.1	27064.0	0.0	0.0	15.4	40.0	20.0	33.3	0.0	0.0
DEX0452_044.nt.2	27063.0	10.0	16.7	26.9	46.7	30.0	42.9	12.5	28.6
DEX0452_044.nt.2	27064.0	0.0	0.0	15.4	40.0	20.0	33.3	0.0	0.0
DEX0452_045.nt.1	30175.0	30.0	30.0	46.2	57.1	60.0	63.2	18.8	25.0

DEX0452_045.nt.1	30176.0	60.0	66.7	46.2	66.7	60.0	75.0	37.5	54.5
DEX0452_046.nt.1	20370.0	0.0	0.0	3.8	4.0	0.0	0.0	6.2	6.2
DEX0452_046.nt.2	20369.0	0.0	0.0	7.7	7.7	5.0	5.0	6.2	6.2
DEX0452_046.nt.2	20370.0	0.0	0.0	3.8	4.0	0.0	0.0	6.2	6.2
DEX0452_047.nt.1	34092.0	30.0	33.3	23.1	23.1	10.0	10.5	43.8	43.8
DEX0452_048.nt.1	26236.0	30.0	30.0	15.4	16.0	15.0	15.8	25.0	25.0
DEX0452_048.nt.1	26237.0	10.0	14.3	15.4	16.0	15.0	15.8	12.5	15.4
DEX0452_049.nt.1	40305.0	10.0	11.1	11.5	12.5	20.0	21.1	0.0	0.0
DEX0452_049.nt.1	40306.0	30.0	30.0	34.6	34.6	50.0	50.0	12.5	12.5
DEX0452_049.nt.2	40305.0	10.0	11.1	11.5	12.5	20.0	21.1	0.0	0.0
DEX0452_049.nt.2	40306.0	30.0	30.0	34.6	34.6	50.0	50.0	12.5	12.5
DEX0452_050.nt.1	19465.0	30.0	30.0	46.2	46.2	70.0	70.0	6.2	6.2
DEX0452_052.nt.1	29054.0	0.0	0.0	19.2	20.8	20.0	21.1	6.2	6.7
DEX0452_053.nt.1	41778.0	10.0	10.0	7.7	8.3	15.0	15.8	0.0	0.0
DEX0452_054.nt.1	27617.0	10.0	10.0	19.2	19.2	25.0	25.0	6.2	6.2
DEX0452_054.nt.1	27618.0	10.0	10.0	15.4	15.4	20.0	20.0	6.2	6.2
DEX0452_055.nt.1	22448.0	10.0	10.0	30.8	30.8	25.0	25.0	25.0	25.0
DEX0452_056.nt.1	14317.0	10.0	10.0	34.6	34.6	20.0	20.0	37.5	37.5
DEX0452_056.nt.1	15115.0	10.0	25.0	0.0	0.0	5.0	9.1	0.0	0.0
DEX0452_056.nt.1	26101.0	10.0	10.0	26.9	26.9	15.0	15.0	31.2	31.2
DEX0452_057.nt.1	24447.0	10.0	10.0	26.9	26.9	30.0	30.0	12.5	12.5
DEX0452_058.nt.2	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.2	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.3	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.3	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.4	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.4	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.5	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.5	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.6	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.6	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.7	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.7	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.8	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.8	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2
DEX0452_058.nt.9	30041.0	10.0	10.0	50.0	50.0	50.0	50.0	25.0	25.0
DEX0452_058.nt.9	30042.0	0.0	0.0	34.6	34.6	40.0	40.0	6.2	6.2

Table 3.

DEX ID	Oligo Name	Mam Multi-Cancer ALL %up n=20	Mam Multi-Cancer ALL %valid up n=20	Mam Multi-Cancer ST1 %up n=9	Mam Multi-Cancer ST1 %valid up n=9	Mam Multi-Cancer ST2,3 %up n=11	Mam Multi-Cancer ST2,3 %valid up n=11
DEX0452_012.nt.1	96143.1	15.0	15.8	0.0	0.0	27.3	27.3
DEX0452_012.nt.1	96144.0	15.0	15.0	0.0	0.0	27.3	27.3
DEX0452_012.nt.1	96144.1	10.0	10.0	0.0	0.0	18.2	18.2
DEX0452_042.nt.1	1689.0	35.0	35.0	44.4	44.4	27.3	27.3

COLON CANCER CHIPS

5 For colon cancer two different chip designs were evaluated with overlapping sets of a total of 38 samples, comparing the expression patterns of colon cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 7 normal colon tissues. For the

Colon Array Chip all 38 samples (23 Ascending colon carcinomas and 15 Rectosigmoidal carcinomas including: 5 stage I cancers, 15 stage II cancers, 15 stage III and 2 stage IV cancers, as well as 28 Grade 1/2 and 10 Grade 3 cancers) were analyzed. The histopathologic grades for cancer are classified as follows: GX, cannot be assessed; G1, 5 well differentiated; G2, Moderately differentiated; G3, poorly differentiated; and G4, undifferentiated. AJCC Cancer Staging Handbook, 5th Edition, 1998, page 9. For the Colon Array Chip analysis, samples were further divided into groups based on the expression pattern of the known colon cancer associated gene Thymidilate Synthase (TS) (13 TS up 25 TS not up). The association of TS with advanced colorectal cancer is well 10 documented. Paradiso *et al.*, *Br J Cancer* 82(3):560-7 (2000); Etienne *et al.*, *J. Clin Oncol.* 20(12):2832-43 (2002); Aschele *et al.* *Clin Cancer Res.* 6(12):4797-802 (2000). For the Multi-Cancer Array Chip a subset of 27 of these samples (14 Ascending colon carcinomas and 13 Rectosigmoidal carcinomas including: 3 stage I cancers, 9 stage II cancers, 13 stage III and 2 stage IV cancers) were assessed.

15 The results for the statistically significant up-regulated genes on the Colon Array
Chip are shown in Table 4 and 5. The results for the statistically significant up-regulated
genes on the Multi-Cancer Array Chip are shown in Table 6.

The first two columns of each table contain information about the sequence itself (Seq ID, Oligo Name), the next columns show the results obtained for all (“ALL”) the colon samples, ascending colon carcinomas (“ASC”), Rectosigmoidal carcinomas (“RS”), cancers corresponding to stages I and II (“ST1,2”), stages III and IV (“ST3,4”), grades 1 and 2 (“GR1,2”), grade 3 (“GR3”), cancers exhibiting up-regulation of the TS gene (“TSup”) or those not exhibiting up-regulation of the TS gene (“NOT TSup”). ‘%up’ indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed n=38 for the Colon Array Chip (n=27 for the Multi-Cancer Array Chip), ‘%valid up’ indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 4.

DEX0452_004.nt.1	40032.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_007.nt.1	28637.0	73.7	73.7	82.6	82.6	60.0	60.0	70.0	70.0	77.8	77.8	
DEX0452_007.nt.1	28638.0	65.8	65.8	73.9	73.9	53.3	53.3	70.0	70.0	61.1	61.1	
DEX0452_024.nt.1	35460.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_024.nt.1	35461.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_040.nt.1	28517.0	5.3	5.3	8.7	8.7	0.0	0.0	5.0	5.0	5.6	5.6	
DEX0452_040.nt.1	28518.0	5.3	5.3	8.7	8.7	0.0	0.0	5.0	5.0	5.6	5.6	
DEX0452_041.nt.1	32006.0	2.6	2.6	4.3	4.3	0.0	0.0	5.0	5.0	0.0	0.0	

Table 5.

DEX ID	Oligo Name	Cln GR1, %up n=28	Cln GR1,2 %valid up n=28	Cln GR3 %up n=10	Cln GR3 %valid up n=10	Cln TS up %up n=13	Cln TS up %valid n=13	Cln NOT TS up %up n=25	Cln NOT TS up %valid n=25
DEX0452_004.nt.1	40031.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_004.nt.1	40032.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_007.nt.1	28637.0	71.4	71.4	80.0	80.0	69.2	69.2	76.0	76.0
DEX0452_007.nt.1	28638.0	64.3	64.3	70.0	70.0	61.5	61.5	68.0	68.0
DEX0452_024.nt.1	35460.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_024.nt.1	35461.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_040.nt.1	28517.0	3.6	3.6	10.0	10.0	7.7	7.7	4.0	4.0
DEX0452_040.nt.1	28518.0	3.6	3.6	10.0	10.0	7.7	7.7	4.0	4.0
DEX0452_041.nt.1	32006.0	3.6	3.6	0.0	0.0	0.0	0.0	4.0	4.0

Table 6.

DEX ID	Oligo Name	Cln Multi-Cancer ALL %up n=27	Cln Multi-Cancer ALL %valid up n=27	Cln Multi-Cancer ASC %up n=14	Cln Multi-Cancer ASC %valid up n=14	Cln Multi-Cancer RS %up n=13	Cln Multi-Cancer RS %valid up n=13
DEX0452_012.nt.1	96143.1	11.1	11.1	14.3	14.3	7.7	7.7
DEX0452_012.nt.1	96144.0	7.4	7.4	14.3	14.3	0.0	0.0
DEX0452_012.nt.1	96144.1	11.1	11.1	14.3	14.3	7.7	7.7
DEX0452_042.nt.1	1689.0	3.7	3.7	7.1	7.1	0.0	0.0

5

LUNG CANCER CHIPS

- For lung cancer two different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of lung cancer derived polyA+ RNA to polyA+ RNA isolated from a pool of 12 normal lung tissues. For the Lung Array Chip all 29 samples (15 squamous cell carcinomas and 14 adenocarcinomas including 14 stage I and 15 stage II/III cancers) were analyzed and for the Multi-Cancer Array Chip a

subset of 22 of these samples (10 squamous cell carcinomas, 12 adenocarcinomas) were assessed.

The results for the statistically significant up-regulated genes on the Lung Array Chip are shown in Table 7. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 8. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next 5 columns show the results obtained for all ("ALL") lung cancer samples, squamous cell carcinomas ("SQ"), adenocarcinomas ("AD"), or cancers corresponding to stage I ("ST1"), or stages II and III ("ST2,3"). '%up' indicates the percentage of all experiments 10 in which up-regulation of at least 2-fold was observed (n=29 for Lung Array Chip, n=22 for Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 7.

DEX ID	Oligo Name	Lng ALL %up n=29	Lng ALL %valid up n=29	Lng SQ %up n=15	Lng SQ %valid up n=15	Lng AD %up n=14	Lng AD %valid up n=14	Lng ST1 %up n=14	Lng ST1 %valid up n=14	Lng ST2,3 %up n=15	Lng ST2,3 %valid up n=15
DEX0452_042.nt.1	1688.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	3540.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	3541.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_043.nt.1	4779.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

15 Table 8.

DEX ID	Oligo Name	Lng Multi-Cancer ALL %up n=22	Lng Multi-Cancer ALL %valid up n=22	Lng Multi-Cancer SQ %up n=10	Lng Multi-Cancer SQ %valid up n=10	Lng Multi-Cancer AD %up n=12	Lng Multi-Cancer AD %valid up n=12
DEX0452_012.nt.1	96143.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	1689.0	0.0	0.0	0.0	0.0	0.0	0.0

OVARIAN CANCER CHIPS

For ovarian cancer two different chip designs were evaluated with overlapping sets of a total of 19 samples, comparing the expression patterns of ovarian cancer derived total RNA to total RNA isolated from a pool of 9 normal ovarian tissues. For the Multi-Cancer Array Chip, all 19 samples (14 invasive carcinomas, 5 low malignant potential samples
5) were analyzed and for the Ovarian Array Chip, a subset of 17 of these samples (13 invasive carcinomas, 4 low malignant potential samples) were assessed.

The results for the statistically significant up-regulated genes on the Ovarian Array Chip are shown in Table 9. The results for the Multi-Cancer Array Chip are shown in Table 10. The first two columns of each table contain information about the sequence 10 itself (DEX ID, Oligo Name), the next columns show the results obtained for all ("ALL") ovarian cancer samples, invasive carcinomas ("INV") and low malignant potential ("LMP") samples. "%up" indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=19 for the Multi-Cancer Array Chip, n=17 for the Ovarian Array Chip), "%valid up" indicates the percentage of experiments with 15 valid expression values in which up-regulation of at least 2-fold was observed.

Table 9.

DEX ID	Oligo Name	Ovr ALL %up n=17	Ovr ALL %valid up n=17	Ovr INV %up n=13	Ovr INV %valid up n=13	Ovr LMP %up n=4	Ovr LMP %valid up n=4
DEX0452_004.nt.1	12147.01	5.9	5.9	7.7	7.7	0.0	0.0
DEX0452_004.nt.1	12147.02	5.9	5.9	7.7	7.7	0.0	0.0
DEX0452_004.nt.1	16301.01	5.9	5.9	7.7	7.7	0.0	0.0
DEX0452_004.nt.1	16301.02	5.9	5.9	7.7	7.7	0.0	0.0
DEX0452_053.nt.1	15931.01	5.9	5.9	7.7	7.7	0.0	0.0
DEX0452_053.nt.1	15931.02	5.9	5.9	7.7	7.7	0.0	0.0

Table 10.

DEX ID	Oligo Name	Ovr Multi-Cancer ALL %up n=19	Ovr Multi-Cancer ALL %valid up n=19	Ovr Multi-Cancer INV %up n=14	Ovr Multi-Cancer INV %valid up n=14	Ovr Multi-Cancer LMP %up n=5	Ovr Multi-Cancer LMP %valid up n=5
DEX0452_012.nt.1	96143.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.0	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.1	0.0	0.0	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	1689.0	21.1	21.1	21.4	21.4	20.0	20.0

20 PROSTATE CANCER

For prostate cancer three different chip designs were evaluated with overlapping sets of a total of 29 samples, comparing the expression patterns of prostate cancer or

benign disease derived total RNA to total RNA isolated from a pool of 35 normal prostate tissues. For the Prostate1 Array and Prostate2 Array Chips all 29 samples (17 prostate cancer samples, 12 non-malignant disease samples) were analyzed. For the Multi-Cancer Array Chip a subset of 28 of these samples (16 prostate cancer samples, 12 non-malignant disease samples) was analyzed.

The results for the statistically significant up-regulated genes on the Prostate1 Array Chip and the Prostate2 Array Chip are shown in Table 11. The results for the statistically significant up-regulated genes on the Multi-Cancer Array Chip are shown in Table 12. The first two columns of each table contain information about the sequence itself (DEX ID, Oligo Name), the next columns show the results obtained for prostate cancer samples ("CAN") or non-malignant disease samples ("DIS"). '%up' indicates the percentage of all experiments in which up-regulation of at least 2-fold was observed (n=29 for the Prostate2 Array Chip and the Multi-Cancer Array Chip), '%valid up' indicates the percentage of experiments with valid expression values in which up-regulation of at least 2-fold was observed.

Table 11.

DEX ID	Oligo Name	Pro CAN %up n=17	Pro CAN %valid up n=17	Pro DIS %up n=12	Pro DIS %valid up n=12
DEX0452_013.nt.2	27919.01	0.0	0.0	0.0	0.0
DEX0452_013.nt.2	27919.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	34478.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	34478.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	34478.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35642.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35642.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35642.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35662.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35662.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.1	35662.03	5.9	5.9	0.0	0.0
DEX0452_015.nt.2	34478.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	34478.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	34478.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35642.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35642.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35642.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35662.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35662.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.2	35662.03	5.9	5.9	0.0	0.0
DEX0452_015.nt.3	34478.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	34478.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	34478.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	35642.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	35642.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	35642.03	0.0	0.0	0.0	0.0

DEX0452_015.nt.3	35662.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	35662.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.3	35662.03	5.9	5.9	0.0	0.0
DEX0452_015.nt.4	34478.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	34478.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	34478.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35642.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35642.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35642.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35662.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35662.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.4	35662.03	5.9	5.9	0.0	0.0
DEX0452_015.nt.5	34478.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	34478.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	34478.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35642.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35642.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35642.03	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35662.01	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35662.02	0.0	0.0	0.0	0.0
DEX0452_015.nt.5	35662.03	5.9	5.9	0.0	0.0
DEX0452_020.nt.1	23434.01	5.9	5.9	0.0	0.0
DEX0452_020.nt.1	23434.02	11.8	11.8	0.0	0.0
DEX0452_020.nt.1	23438.01	5.9	5.9	0.0	0.0
DEX0452_020.nt.1	23438.02	0.0	0.0	0.0	0.0
DEX0452_020.nt.1	23482.01	11.8	11.8	0.0	0.0
DEX0452_020.nt.1	23482.02	11.8	11.8	0.0	0.0
DEX0452_020.nt.1	23536.01	5.9	6.2	0.0	0.0
DEX0452_020.nt.1	23536.02	5.9	5.9	0.0	0.0
DEX0452_020.nt.1	27967.01	5.9	5.9	0.0	0.0
DEX0452_020.nt.1	27967.02	11.8	11.8	0.0	0.0
DEX0452_043.nt.1	34916.01	0.0	0.0	0.0	0.0
DEX0452_043.nt.1	34916.02	0.0	0.0	0.0	0.0

Table 12.

DEX ID	Oligo Name	Pro Multi-Cancer CAN % ^{up} n=16	Pro Multi-Cancer CAN %valid up n=16	Pro Multi-Cancer DIS % ^{up} n=12	Pro Multi-Cancer DIS %valid up n=12
DEX0452_012.nt.1	96143.1	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.0	0.0	0.0	0.0	0.0
DEX0452_012.nt.1	96144.1	0.0	0.0	0.0	0.0
DEX0452_042.nt.1	1689.0	0.0	0.0	0.0	0.0

SEQ ID NO: 1-95 was up-regulated on various tissue microarrays. Accordingly,

- 5 nucleotide SEQ ID NO: 1-95 or the encoded protein SEQ ID NO: 96-232 may be used as a cancer therapeutic and/or diagnostic target for the tissues in which expression is shown.

The following table lists the location (Oligo Location) where the microarray oligos (Oligo ID) map on the transcripts (DEX ID) of the present invention. Each Oligo ID may 10 have been printed multiple times on a single chip as replicates. The Oligo Name is an

exemplary replicate (e.g. 1000.01) for the Oligo ID (e.g. 1000), and data from other replicates (e.g. 1000.02, 1000.03) may be reported. Additionally, the Array (Chip Name) that each oligo and oligo replicates were printed on is included.

DEX NT ID	Oligo ID	Oligo Name	Chip Name	Oligo Location
DEX0452_001.nt.1	34133	34133.0	Breast array	2774-2833
DEX0452_001.nt.1	34132	34132.0	Breast array	3024-3083
DEX0452_002.nt.1	13283	13283.0	Breast array	1216-1275
DEX0452_002.nt.1	13284	13284.0	Breast array	1176-1235
DEX0452_003.nt.1	14380	14380.0	Breast array	1388-1447
DEX0452_003.nt.1	14381	14381.0	Breast array	1338-1397
DEX0452_003.nt.2	14380	14380.0	Breast array	775-834
DEX0452_003.nt.2	14381	14381.0	Breast array	725-784
DEX0452_004.nt.1	12147	12147.01	Ovarian array	380-439
DEX0452_004.nt.1	40031	40031.0	Colon array	1477-1536
DEX0452_004.nt.1	40032	40032.0	Colon array	1447-1506
DEX0452_004.nt.1	16301	16301.02	Ovarian array	298-357
DEX0452_004.nt.1	28910	28910.0	Breast array	256-315
DEX0452_005.nt.1	16289	16289.0	Breast array	637-696
DEX0452_005.nt.1	29727	29727.0	Breast array	1585-1644
DEX0452_005.nt.1	29728	29728.0	Breast array	1327-1386
DEX0452_005.nt.1	16290	16290.0	Breast array	543-602
DEX0452_006.nt.1	20370	20370.0	Breast array	2786-2845
DEX0452_006.nt.1	20369	20369.0	Breast array	2955-3014
DEX0452_007.nt.1	28637	28637.0	Colon array	715-774
DEX0452_007.nt.1	12616	12616.0	Breast array	515-574
DEX0452_007.nt.1	12615	12615.0	Breast array	535-594
DEX0452_007.nt.1	28638	28638.0	Colon array	575-634
DEX0452_008.nt.1	27530	27530.0	Breast array	1115-1174
DEX0452_009.nt.1	20207	20207.0	Breast array	151-210
DEX0452_009.nt.2	20208	20208.0	Breast array	1158-1217
DEX0452_009.nt.2	20207	20207.0	Breast array	1229-1288
DEX0452_010.nt.1	31614	31614.0	Breast array	2198-2257
DEX0452_010.nt.1	15032	15032.0	Breast array	1164-1223
DEX0452_010.nt.1	15033	15033.0	Breast array	1065-1124
DEX0452_010.nt.1	31615	31615.0	Breast array	2114-2173
DEX0452_011.nt.1	31927	31927.0	Breast array	513-572
DEX0452_012.nt.1	96144	96144.0	Multi-Cancer array	5222-5281
DEX0452_012.nt.1	96143	96143.0	Multi-Cancer array	5262-5321
DEX0452_013.nt.1	11156	11156.0	Breast array	2780-2839
DEX0452_013.nt.2	27919	27919.02	Prostate1 array	4453-4512
DEX0452_014.nt.1	38922	38922.0	Breast array	467-526
DEX0452_014.nt.1	38921	38921.0	Breast array	598-657
DEX0452_015.nt.1	34478	34478.02	Prostate2 array	1797-1856
DEX0452_015.nt.1	18118	18118.0	Breast array	1797-1856

DEX0452_015.nt.1	18256	18256.0	Breast array	1797-1856
DEX0452_015.nt.1	18250	18250.0	Breast array	1991-2050
DEX0452_015.nt.1	35642	35642.03	Prostate2 array	1797-1856
DEX0452_015.nt.1	35662	35662.03	Prostate2 array	1991-2050
DEX0452_015.nt.2	18250	18250.0	Breast array	1356-1415
DEX0452_015.nt.2	34478	34478.02	Prostate2 array	1162-1221
DEX0452_015.nt.2	35662	35662.03	Prostate2 array	1356-1415
DEX0452_015.nt.2	18118	18118.0	Breast array	1162-1221
DEX0452_015.nt.2	35642	35642.03	Prostate2 array	1162-1221
DEX0452_015.nt.2	18256	18256.0	Breast array	1162-1221
DEX0452_015.nt.3	35662	35662.03	Prostate2 array	1193-1252
DEX0452_015.nt.3	34478	34478.02	Prostate2 array	999-1058
DEX0452_015.nt.3	18250	18250.0	Breast array	1193-1252
DEX0452_015.nt.3	18256	18256.0	Breast array	999-1058
DEX0452_015.nt.3	35642	35642.03	Prostate2 array	999-1058
DEX0452_015.nt.3	18118	18118.0	Breast array	999-1058
DEX0452_015.nt.4	18256	18256.0	Breast array	532-591
DEX0452_015.nt.4	35642	35642.03	Prostate2 array	532-591
DEX0452_015.nt.4	34478	34478.02	Prostate2 array	532-591
DEX0452_015.nt.4	18250	18250.0	Breast array	726-785
DEX0452_015.nt.4	35662	35662.03	Prostate2 array	726-785
DEX0452_015.nt.4	18118	18118.0	Breast array	532-591
DEX0452_015.nt.5	18256	18256.0	Breast array	337-396
DEX0452_015.nt.5	35642	35642.03	Prostate2 array	337-396
DEX0452_015.nt.5	18118	18118.0	Breast array	337-396
DEX0452_015.nt.5	35662	35662.03	Prostate2 array	531-590
DEX0452_015.nt.5	18250	18250.0	Breast array	531-590
DEX0452_015.nt.5	34478	34478.02	Prostate2 array	337-396
DEX0452_016.nt.1	40284	40284.0	Breast array	3156-3215
DEX0452_016.nt.1	40273	40273.0	Breast array	3227-3286
DEX0452_016.nt.1	19496	19496.0	Breast array	3168-3227
DEX0452_016.nt.2	20286	20286.0	Breast array	3347-3406
DEX0452_016.nt.2	20285	20285.0	Breast array	3390-3449
DEX0452_016.nt.2	19496	19496.0	Breast array	3809-3868
DEX0452_016.nt.2	40273	40273.0	Breast array	3868-3927
DEX0452_016.nt.2	40284	40284.0	Breast array	3797-3856
DEX0452_016.nt.3	40273	40273.0	Breast array	3908-3967
DEX0452_016.nt.3	19496	19496.0	Breast array	3849-3908
DEX0452_016.nt.3	40284	40284.0	Breast array	3837-3896
DEX0452_016.nt.4	19496	19496.0	Breast array	3366-3425
DEX0452_016.nt.4	40284	40284.0	Breast array	3354-3413
DEX0452_016.nt.4	40273	40273.0	Breast array	3425-3484
DEX0452_016.nt.5	19497	19497.0	Breast array	4785-4844
DEX0452_016.nt.5	20285	20285.0	Breast array	3878-3937
DEX0452_016.nt.5	20286	20286.0	Breast array	3835-3894
DEX0452_016.nt.5	40284	40284.0	Breast array	4813-4872
DEX0452_016.nt.5	40273	40273.0	Breast array	4884-4943

DEX0452_016.nt.5	19496	19496.0	Breast array	4825-4884
DEX0452_016.nt.6	20286	20286.0	Breast array	3835-3894
DEX0452_016.nt.6	40273	40273.0	Breast array	4492-4551
DEX0452_016.nt.6	20285	20285.0	Breast array	3878-3937
DEX0452_016.nt.6	40284	40284.0	Breast array	4421-4480
DEX0452_016.nt.6	19496	19496.0	Breast array	4433-4492
DEX0452_017.nt.1	25675	25675.0	Breast array	1030-1089
DEX0452_017.nt.1	25674	25674.0	Breast array	1101-1160
DEX0452_018.nt.1	21562	21562.0	Breast array	5986-6045
DEX0452_018.nt.1	21561	21561.0	Breast array	6216-6275
DEX0452_019.nt.1	12954	12954.0	Breast array	1142-1201
DEX0452_019.nt.1	12953	12953.0	Breast array	1204-1263
DEX0452_020.nt.1	17938	17938.0	Breast array	752-811
DEX0452_020.nt.1	27967	27967.02	Prostatael array	752-811
DEX0452_020.nt.1	23536	23536.02	Prostatael array	1059-1118
DEX0452_020.nt.1	17933	17933.0	Breast array	958-1017
DEX0452_020.nt.1	23482	23482.02	Prostatael array	753-812
DEX0452_020.nt.1	17934	17934.0	Breast array	567-626
DEX0452_020.nt.1	23438	23438.02	Prostatael array	567-626
DEX0452_020.nt.1	23434	23434.01	Prostatael array	752-811
DEX0452_020.nt.1	17942	17942.0	Breast array	753-812
DEX0452_020.nt.1	17932	17932.0	Breast array	1059-1118
DEX0452_021.nt.1	25824	25824.0	Breast array	2318-2377
DEX0452_022.nt.1	29794	29794.0	Breast array	154-213
DEX0452_022.nt.1	29793	29793.0	Breast array	388-447
DEX0452_023.nt.1	19175	19175.0	Breast array	1258-1317
DEX0452_023.nt.1	19174	19174.0	Breast array	1281-1340
DEX0452_024.nt.1	13892	13892.0	Breast array	277-336
DEX0452_024.nt.1	35461	35461.0	Colon array	536-595
DEX0452_024.nt.1	35460	35460.0	Colon array	576-635
DEX0452_025.nt.1	18383	18383.0	Breast array	500-559
DEX0452_026.nt.1	35953	35953.0	Breast array	902-961
DEX0452_026.nt.1	35954	35954.0	Breast array	812-871
DEX0452_027.nt.1	33040	33040.0	Breast array	1983-2042
DEX0452_027.nt.1	33041	33041.0	Breast array	1795-1854
DEX0452_027.nt.2	33040	33040.0	Breast array	1228-1287
DEX0452_027.nt.2	33041	33041.0	Breast array	1040-1099
DEX0452_029.nt.1	19254	19254.0	Breast array	1349-1408
DEX0452_029.nt.1	33276	33276.0	Breast array	2849-2908
DEX0452_029.nt.1	19255	19255.0	Breast array	1325-1384
DEX0452_029.nt.1	33277	33277.0	Breast array	2809-2868
DEX0452_029.nt.2	33276	33276.0	Breast array	922-981
DEX0452_029.nt.2	33277	33277.0	Breast array	882-941
DEX0452_030.nt.1	27825	27825.0	Breast array	498-557
DEX0452_030.nt.1	27826	27826.0	Breast array	344-403
DEX0452_031.nt.1	32497	32497.0	Breast array	511-570
DEX0452_031.nt.1	32496	32496.0	Breast array	552-611

DEX0452_031.nt.2	32497	32497.0	Breast array	511-570
DEX0452_031.nt.2	32496	32496.0	Breast array	552-611
DEX0452_031.nt.3	32497	32497.0	Breast array	511-570
DEX0452_031.nt.3	32496	32496.0	Breast array	552-611
DEX0452_032.nt.1	40320	40320.0	Breast array	506-565
DEX0452_032.nt.1	31576	31576.0	Breast array	943-1002
DEX0452_032.nt.1	31577	31577.0	Breast array	899-958
DEX0452_032.nt.1	40363	40363.0	Breast array	444-503
DEX0452_032.nt.1	40364	40364.0	Breast array	404-463
DEX0452_034.nt.1	25930	25930.0	Breast array	807-866
DEX0452_034.nt.1	25931	25931.0	Breast array	787-846
DEX0452_034.nt.2	25930	25930.0	Breast array	999-1058
DEX0452_034.nt.3	25931	25931.0	Breast array	866-925
DEX0452_035.nt.1	27220	27220.0	Breast array	1532-1591
DEX0452_036.nt.1	27219	27219.0	Breast array	2237-2296
DEX0452_036.nt.2	27220	27220.0	Breast array	2424-2483
DEX0452_036.nt.2	27219	27219.0	Breast array	2464-2523
DEX0452_037.nt.1	27234	27234.0	Breast array	2317-2376
DEX0452_037.nt.1	40267	40267.0	Breast array	836-895
DEX0452_037.nt.1	27233	27233.0	Breast array	2358-2417
DEX0452_037.nt.2	27234	27234.0	Breast array	1030-1089
DEX0452_037.nt.2	27233	27233.0	Breast array	1071-1130
DEX0452_038.nt.1	40103	40103.0	Breast array	1363-1422
DEX0452_039.nt.1	27218	27218.0	Breast array	523-582
DEX0452_039.nt.1	12621	12621.0	Breast array	268-327
DEX0452_039.nt.1	12631	12631.0	Breast array	523-582
DEX0452_039.nt.1	12622	12622.0	Breast array	181-240
DEX0452_039.nt.1	27217	27217.0	Breast array	886-945
DEX0452_040.nt.1	28517	28517.0	Colon array	441-500
DEX0452_040.nt.1	28518	28518.0	Colon array	213-272
DEX0452_040.nt.1	24443	24443.0	Breast array	348-407
DEX0452_040.nt.1	24442	24442.0	Breast array	441-500
DEX0452_041.nt.1	20612	20612.0	Breast array	487-546
DEX0452_041.nt.1	32006	32006.0	Colon array	487-546
DEX0452_042.nt.1	1689	1689.0	Multi-Cancer array	4181-4240
DEX0452_042.nt.1	3541	3541.0	Lung array	2988-3047
DEX0452_042.nt.1	27230	27230.0	Breast array	2503-2562
DEX0452_042.nt.1	3540	3540.0	Lung array	2998-3057
DEX0452_042.nt.1	1688	1688.0	Lung array	4183-4242
DEX0452_042.nt.1	27229	27229.0	Breast array	2533-2592
DEX0452_043.nt.1	28899	28899.0	Breast array	56-115
DEX0452_043.nt.1	34916	34916.01	Prostate1 array	56-115
DEX0452_043.nt.1	4779	4779.0	Lung array	56-115
DEX0452_044.nt.1	27063	27063.0	Breast array	1707-1766
DEX0452_044.nt.1	27064	27064.0	Breast array	1488-1547
DEX0452_044.nt.2	27063	27063.0	Breast array	1415-1474

DEX0452_044.nt.2	27064	27064.0	Breast array	1196-1255
DEX0452_045.nt.1	30176	30176.0	Breast array	596-655
DEX0452_045.nt.1	30175	30175.0	Breast array	644-703
DEX0452_046.nt.1	20370	20370.0	Breast array	6933-6992
DEX0452_046.nt.2	20369	20369.0	Breast array	6746-6805
DEX0452_047.nt.1	34092	34092.0	Breast array	2471-2530
DEX0452_048.nt.1	26237	26237.0	Breast array	3124-3183
DEX0452_048.nt.1	26236	26236.0	Breast array	3164-3223
DEX0452_049.nt.1	40305	40305.0	Breast array	467-526
DEX0452_049.nt.1	40306	40306.0	Breast array	368-427
DEX0452_049.nt.2	40305	40305.0	Breast array	382-441
DEX0452_049.nt.2	40306	40306.0	Breast array	283-342
DEX0452_050.nt.1	19465	19465.0	Breast array	24-83
DEX0452_052.nt.1	29054	29054.0	Breast array	1375-1434
DEX0452_053.nt.1	15931	15931.02	Ovarian array	469-528
DEX0452_053.nt.1	41778	41778.0	Breast array	337-396
DEX0452_054.nt.1	27618	27618.0	Breast array	2791-2850
DEX0452_054.nt.1	27617	27617.0	Breast array	2909-2968
DEX0452_055.nt.1	22448	22448.0	Breast array	4760-4819
DEX0452_056.nt.1	15115	15115.0	Breast array	3963-4022
DEX0452_056.nt.1	26101	26101.0	Breast array	2878-2937
DEX0452_056.nt.1	14317	14317.0	Breast array	2136-2195
DEX0452_057.nt.1	24447	24447.0	Breast array	2768-2827
DEX0452_058.nt.2	30041	30041.0	Breast array	1393-1452
DEX0452_058.nt.2	30042	30042.0	Breast array	1353-1412
DEX0452_058.nt.3	30041	30041.0	Breast array	875-934
DEX0452_058.nt.3	30042	30042.0	Breast array	835-894
DEX0452_058.nt.4	30041	30041.0	Breast array	748-807
DEX0452_058.nt.4	30042	30042.0	Breast array	708-767
DEX0452_058.nt.5	30041	30041.0	Breast array	583-642
DEX0452_058.nt.5	30042	30042.0	Breast array	543-602
DEX0452_058.nt.6	30041	30041.0	Breast array	724-783
DEX0452_058.nt.6	30042	30042.0	Breast array	684-743
DEX0452_058.nt.7	30041	30041.0	Breast array	759-818
DEX0452_058.nt.7	30042	30042.0	Breast array	719-778
DEX0452_058.nt.8	30041	30041.0	Breast array	380-439
DEX0452_058.nt.9	30042	30042.0	Breast array	215-274
DEX0452_058.nt.9	30041	30041.0	Breast array	255-314

Example 2b: Relative Quantitation of Gene Expression

Real-Time quantitative PCR with fluorescent Taqman® probes is a quantitation detection system utilizing the 5'- 3' nuclease activity of Taq DNA polymerase. The 5 method uses an internal fluorescent oligonucleotide probe (Taqman®) labeled with a 5' reporter dye and a downstream, 3' quencher dye. During PCR, the 5'-3' nuclease activity

of Taq DNA polymerase releases the reporter, whose fluorescence can then be detected by the laser detector of the Model 7700 Sequence Detection System (PE Applied Biosystems, Foster City, CA, USA). Amplification of an endogenous control is used to standardize the amount of sample RNA added to the reaction and normalize for Reverse Transcriptase (RT) efficiency. Either cyclophilin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), ATPase, or 18S ribosomal RNA (rRNA) is used as this endogenous control. To calculate relative quantitation between all the samples studied, the target RNA levels for one sample were used as the basis for comparative results (calibrator). Quantitation relative to the "calibrator" can be obtained using the comparative method (User Bulletin 10 #2: ABI PRISM 7700 Sequence Detection System).

The tissue distribution and the level of the target gene are evaluated for every sample in normal and cancer tissues. Total RNA is extracted from normal tissues, cancer tissues, and from cancers and the corresponding matched adjacent tissues. Subsequently, first strand cDNA is prepared with reverse transcriptase and the polymerase chain reaction 15 is done using primers and Taqman® probes specific to each target gene. The results are analyzed using the ABI PRISM 7700 Sequence Detector. The absolute numbers are relative levels of expression of the target gene in a particular tissue compared to the calibrator tissue.

One of ordinary skill can design appropriate primers. The relative levels of 20 expression of the BSNA versus normal tissues and other cancer tissues can then be determined. All the values are compared to the calibrator. Normal RNA samples are commercially available pools, originated by pooling samples of a particular tissue from different individuals.

The relative levels of expression of the BSNA in pairs of matched samples may 25 also be determined. A matched pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. All the values are compared to the calibrator.

In the analysis of matching samples, the BSNA show a high degree of tissue specificity for the tissue of interest. These results confirm the tissue specificity results 30 obtained with normal pooled samples. Further, the level of mRNA expression in cancer samples and the isogenic normal adjacent tissue from the same individual are compared. This comparison provides an indication of specificity for the cancer state (*e.g.* higher levels of mRNA expression in the cancer sample compared to the normal adjacent).

Informaton on the samples tested in the QPCR experiments below include the Sample ID (Smpl ID), Tissue, Tissue Type (Tiss Type), Diagnosis (DIAG), Disease Detail, and Stage or Grade (STG or GRD) in following table.

Sample ID	Tissue	Tissue Type	Diagnosis	Disease Detail	Stage or Grade
355	Mammary	CAN	Invasive lobular carcinoma	Invasive lobular carcinoma	
355	Mammary	NAT	NAT		
B011X	Mammary	CAN		Cancer	
B011X	Mammary	NAT		NAT	
S621	Mammary	CAN	Infiltrating ductal carcinoma	Infiltrating Duct Adenocarcinoma	G3; T1NxMx
S621	Mammary	NAT		NAT	
S516	Mammary	CAN	Infiltrating ductal carcinoma	Infiltrating Ductal Carcinoma with Lymphatic Invasion	Stage I G2; T1NoMo
S516	Mammary	NAT		NAT	
522	Mammary	CAN	Infiltrating ductal carcinoma	Infiltrating ductal carcinoma	G III
522	Mammary	NAT		NAT	
76DN	Mammary	CAN		Invasive ductal carcinoma	G3, poorly diff.
76DN	Mammary	NAT		NAT	
19DN	Mammary	CAN	Invasive ductal carcinoma	Invasive ductal carcinoma	G3, Stage IIA; T2N0M0
19DN	Mammary	NAT		NAT	
42DN	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	T3aN1M0 IIIA, G3
42DN	Mammary	NAT		NAT	
517	Mammary	CAN	Infiltrating ductal carcinoma	Infiltrating ductal carcinoma	St. IIA, G3
517	Mammary	NAT		NAT	
781M	Mammary	CAN	Invasive ductal carcinoma		Architectural grade-3/3, Nuclear grade-3/3
781M	Mammary	NAT		NAT	
869M	Mammary	CAN	Invasive carcinoma	Invasive Carcinoma	Stage IIA G1; T2N0M0
869M	Mammary	NAT		NAT	
976M	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	T2N1M0 (Stage 2B Grade 2-3)
976M	Mammary	NAT		NAT	
S570	Mammary	CAN	Carcinoma	Carcinoma	Stage IIA; T1N1M0
S570	Mammary	NAT		NAT	

S699	Mammary	CAN	Invasive lobular carcinoma	Invasive Lobular Carcinoma	Stage IIB G1; T2N1Mo
S699	Mammary	NAT		NAT	
S997	Mammary	CAN	Invasive ductal carcinoma	Invasive Ductal Carcinoma	Stage IIB G3; T2N1Mo
S997	Mammary	NAT		NAT	
030B	Urinary Bladder	CAN	Carcinoma	invasive Carcinoma, poorly differentiated	Stage III, Grade 3
030B	Urinary Bladder	NAT		NAT	
520B	Urinary Bladder	CAN	Sarcomatoid transitional cell carcinoma	Sarcomatoid transitional cell carcinoma	
520B	Urinary Bladder	NAT		NAT	
TR17	Urinary Bladder	CAN	Carcinoma	Stage II/Grade III	
TR17	Urinary Bladder	NAT		NAT	
401C	Colon	CAN	Adenocarcinoma of ascending colon and cecum	Adenocarcinoma of ascending colon and cecum	Stage III
401C	Colon	NAT		NAT	
AS43	Colon	CAN	Adenocarcinoma	malignant	
AS43	Colon	NAT	Adenocarcinoma	NAT	
AS98	Colon	CAN	Adenocarcinoma	Moderately to poorly differentiated adenocarcinoma	Duke's C
AS98	Colon	NAT		NAT	
CM12	Colon	CAN		T	Stage D
CM12	Colon	NAT	Adenocarcinoma	Nat	
DC19	Colon	CAN		T	Stage B
DC19	Colon	NAT		NL	
RC01	Colon	CAN	Cancer		Stage IV
RC01	Colon	NAT		NAT	
RS53	Colon	CAN	Adenocarcinoma	moderately differentiated adenocarcinoma	
RS53	Colon	NAT	Adenocarcinoma	NAT	
SG27	Colon	CAN		malig	Stage B
SG27	Colon	NAT		NAT	
TX01	Colon	CAN	Adenocarcinoma	Moderately differentiated adenocarcinoma of cecum	Stage II; T3NoMo
TX01	Colon	NAT		NAT	

KS52	Cervix	CAN	Squamous cell carcinoma	Keratinizing Squamous Cell Carcinoma	IIIB, well diff. G1; T3bNxM0
KS52	Cervix	NAT		NAT	
NK23	Cervix	CAN		Nonkeratinizing Large Cell	FIGO IIIB, undiff. G4; T3bNxM0
NK23	Cervix	NAT		NAT	
NKS54	Cervix	CAN	Squamous cell carcinoma	Nonkeratinizing Squamous Cell Carcinoma	IIB, mod diff. G2; T2bNxM0
NKS54	Cervix	NAT		NAT	
NKS55	Cervix	CAN	Squamous cell carcinoma	Nonkeratinizing Squamous Cell Carcinoma	IIIB, Mod diff. G2; T3bNxM0
NKS55	Cervix	NAT		NAT	
NKS81	Cervix	CAN	Squamous cell carcinoma	large cell nonkeratinizing sq carc, IIB, moderately diff	IIB
NKS81	Cervix	NAT		NAT	
10479	Endometrium	CAN		malignant mixed mullerian tumor	T?, Nx, M1
10479	Endometrium	NAT		NAT	
28XA	Endometrium	CAN	Endometrial adenocarcinoma	malignant	II/III
28XA	Endometrium	NAT		NAT	II/III
8XA	Endometrium	CAN	mod. diff, invasive, squamous differentiation, FIGO-II		
8XA	Endometrium	NAT		NAT	
106XD	Kidney	CAN	Renal cell carcinoma, clear cell, localized		3
106XD	Kidney	NAT		NL	
107XD	Kidney	CAN	Renal cell carcinoma	renal cell carcinoma, clear cell, with metastatic	G III
107XD	Kidney	NAT		NL	
109XD	Kidney	CAN		Malignant	G III
109XD	Kidney	NAT		NL	
10XD	Kidney	CAN	Renal cell carcinoma	renal cell carcinoma, clear cell, localized, grade 2-3	3
10XD	Kidney	NAT		NL	
22K	Kidney	CAN	Renal cell carcinoma	Renal cell carcinoma	G2, Mod. Diff.
22K	Kidney	NAT		NAT	
15XA	Liver	CAN		Sarcoma, Retropertitoneal Tumor	Grade-2

15XA	Liver	NAT		CA	St. I, G4
174L	Liver	CAN	Hepatocellular carcinoma	Moderate to well differentiated hepatocellular carcinoma	
174L	Liver	NAT	Hepatocellular carcinoma	NAT	
187L	Liver	CAN	Adenocarcinoma	Metastatic Adenocarcinoma	Liver (Gallbladder)
187L	Liver	NAT		NAT	
205L	Lung	CAN	Adenocarcinoma	poorly differentiated adenocarcinoma	T2, N1, Mx
205L	Lung	NAT		NAT	
315L	Lung	CAN	Squamous cell carcinoma		
315L	Lung	NAT	Adenocarcinoma	NAT	
507L	Lung	CAN	Bronchioloalveolar carcinoma	bronchioalveolar carcinoma	Stage IB, G1, well diff.
507L	Lung	NAT		NAT	
528L	Lung	CAN	Adenocarcinoma	Adenocarcinoma	St.IV, T2N0M1 infiltrating poorly diff.
528L	Lung	NAT		NAT	
8837L	Lung	CAN	Squamous cell carcinoma	Squamous cell carcinoma	T2, N0, M0
8837L	Lung	NAT		NAT	
AC11	Lung	CAN	Adenocarcinoma	poorly differentiated adenocarcinoma	T2, N2, M1
AC11	Lung	NAT		NAT	
AC39	Lung	CAN	Adenocarcinoma	intermediate grade adenocarcinoma	T2, N2, Mx
AC39	Lung	NAT		NAT	
SQ80	Lung	CAN	Squamous cell carcinoma	poorly differentiated squamous cell carcinoma	T1, N1, M0
SQ80	Lung	NAT		NAT	
SQ81	Lung	CAN	Squamous cell carcinoma	poorly differentiated squamous carcinoma	T3, N1, Mx
SQ81	Lung	NAT		NAT	
G021	Ovary	CAN	Carcinoma	St. IIIC, poorly diff.	Stage- IIIC, poorly diff.
G021	Ovary	NAT		NAT	
206I	Ovary	NRM		NL	
5150	Ovary	NRM		Normal	
18GA	Ovary	NRM		NL	
3370	Ovary	NRM		Normal	
1230	Ovary	NRM		Normal	

C177	Ovary	NRM		several fluid filled cysts	
40G	Ovary	NRM		NL	
10050				papillary serous and endometrioid ovarian carcinoma, concurrent metastatic breast cancer	3
10400	Ovary	CAN		papillary serous adeno, metastatic	
1050	Ovary	CAN		Papillary Serous Carcinoma with Focal Mucinous Differentiation	Stage IC G0; T1cN0M0
130X	Ovary	CAN		Ovarian cancer	
C004	Ovary	NRM		NL	
7180	Ovary	CAN	Adenocarcinoma	malignant tumor	IIIC
A1B	Ovary	CAN	Adenocarcinoma	CA	
71XL	Pancreas	CAN		villous adenoma with paneth cell metaplasia	localized
71XL	Pancreas	NAT		NL	
82XP	Pancreas	CAN		serious cystadenoma	
82XP	Pancreas	NAT		NL	
92X	Pancreas	CAN	Ductal adenocarcinoma	ductal adenocarcinoma	mod to focally poorly diff.
92X	Pancreas	NAT		NL	
23B	Prostate	CAN		Prostate tumor	Gleason's 3+4
23B	Prostate	NAT		NAT	
675P	Prostate	CAN	Adenocarcinoma	adenocarcinoma	
675P	Prostate	NAT		Normal	
958P	Prostate	CAN	Adenocarcinoma	Adenocarcinoma	T2C, NO, MX
958P	Prostate	NAT		NAT	
65XB	Prostate	CAN	Adenocarcinoma	adenocarcinoma	3+4=7
65XB	Prostate	NAT		NL	
84XB	Prostate	CAN	Adenocarcinoma	adenocarcinoma	2+3
84XB	Prostate	NAT		NL	
855P	Prostate	BPH		BPH	
276P	Prostate	BPH		BPH	
767B	Prostate	BPH		prostate BPH	
263C	Prostate	BPH		BPH	
10R	Prostate	PRO ST		active chronic prostatitis	T0, NO, MO
20R	Prostate	PRO ST		PROSTATITIS	
39A	Skin	CAN		CA	St. II

39A	Skin	NAT		CA	St. II
287S	Skin	CAN	Squamous cell carcinoma	Invasive Keratinizing Squamous Cell Carcinoma	Moderately Differentiated
287S	Skin	NAT		NAT	
669S	Skin	CAN	Melanoma	Nodular malignant melanoma	
669S	Skin	NAT		NAT	
171S	Small Intestine	CAN	Adenocarcinoma	Moderately differentiated Adenocarcinoma, invasive	
171S	Small Intestine	NAT		NAT	
H89	Small Intestine	CAN	Adenocarcinoma	Adenocarcinoma	80% tumor, 50% necrosis, moderately differentiated, G2-3; T3N1MX
H89	Small Intestine	NAT	Adenocarcinoma	NAT	
20SM	Small Intestine	CAN	Adenocarcinoma	Adenocarcinoma, metastatic to lung & liver	St. IV, poorly diff.
20SM	Small Intestine	NAT		NAT	
88S	Stomach	CAN	Adenocarcinoma	Mucinous adenocarcinoma	T3N1M0, St. IIIA
88S	Stomach	NAT		NAT	
261S	Stomach	CAN	Signet-ring cell carcinoma	Signet-ring cell carcinoma	Stage IIIA, T3N1M0
261S	Stomach	NAT		NAT	
288S	Stomach	CAN	Adenocarcinoma	Infiltrating Adneocarcinoma	Moderately Differentiated
288S	Stomach	NAT		NAT	
AC93 or 509L	Stomach	CAN	Adenocarcinoma	Adenocarcinoma	St. IV, G4, T4N3M0, poorly diff.
AC93 or 509L	Stomach	NAT		NAT	
39X	Testes	CAN		CA	
39X	Testes	NAT		NAT	
647T	Testes	CAN	Teratocarcinoma	Teratocarcinoma	Stage IA
647T	Testes	NAT	Teratocarcinoma	NAT	
663T	Testes	CAN	Teratocarcinoma	Teratocarcinoma	
663T	Testes	NAT		NAT	
56T	Thyroid Gland	CAN	Papillary carcinoma	Papillary Carcinoma	St. III; T4N1M0

56T	Thyroid Gland	NAT		NAT	
143N	Thyroid Gland	CAN	Follicular carcinoma	Follicular Carcinoma	
143N	Thyroid Gland	NAT		NAT	
270T	Thyroid Gland	CAN		CA	
270T	Thyroid Gland	NAT		NAT	
135XO	Uterus	CAN		Uterus normal	
135XO	Uterus	NAT		Uterus tumor	
85XU	Uterus	CAN		endometrial carcinoma	I
85XU	Uterus	NAT		NL	
B1	Blood	NRM		Normal	
B3	Blood	NRM		Normal	
B5	Blood	NRM		Normal	
B6	Blood	NRM		Normal	
B11	Blood	NRM		Normal	
982B	Blood	NRM		Normal	
48AD	Adrenal Gland	NRM		Normal	
10BR	Brain	NRM		Normal	
01CL	Colon	NRM		Normal	
06CV	Cervix	NRM		Normal	
01ES	Esophagus	NRM		Normal	
46HR	Heart	NRM		Normal	
00HR	Human Reference	CAN	CAN	Cancer pool	
55KD	Kidney	NRM		Normal	
89LV	Liver	NRM		Normal	
90LN	Lung	NRM		Normal	
01MA	Mammary	NRM		Normal	
84MU	Skeletal Muscle	NRM		Normal	
3APV	Ovary	NRM		Normal	
04PA	Pancreas	NRM		Normal	
59PL	Placenta	NRM		Normal	
09PR	Prostate	NRM		Normal	
21RC	Rectum	NRM		Normal	
59SM	Small Intestine	NRM		Normal	
7GSP	Spleen	NRM		Normal	
09ST	Stomach	NRM		Normal	
4GTS	Testes	NRM		Normal	
99TM	Thymus Gland	NRM		Normal	
16TR	Trachea	NRM		Normal	
57UT	Uterus	NRM		Normal	

DEX0452_010.nt.1 (Mam113)

The relative expression level of Mam113 in various tissue samples is included below. Tissue samples include 79 pairs of matching samples, 7 non matched cancer samples, and 36 normal samples, all from various tissues annotated in the table. A

matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. Of the normal samples 6 were blood samples which measured the expression levels in blood cells. Additionally, 2 prostatitis, and 4 Benign Prostatic Hyperplasia (BPH) samples are included. All the values are compared to human cancer sample HUMREF00HR (calibrator).

The table below contains the relative expression level values for the sample as compared to the calibrator. The table includes the Sample ID, and expression level values for the following samples: Cancer (CAN), Normal Adjacent Tissue (NAT), Normal Tissue (NRM), Benign Prostatic Hyperplasia (BPH), and Prostatitis (PROST).

Sample ID	CAN	NAT	NRM	BPH	PROST
MAM355	4.58	0.00			
MAMB011X	14.60	6.15			
MAMS621	1.22	0.10			
MAMS516	0.80	0.51			
MAM522	6.52	0.50			
MAM76DN	11.50	0.65			
MAM976M	21.81	0.73			
MAM781M	16.77	0.62			
MAM19DN	19.09	9.18			
MAM517	3.61	1.45			
MAMS997	44.39	7.66			
MAM42DN	24.48	7.51			
MAM869M	5.76	0.23			
MAMS699	3.00	4.37			
MAMS570	3.93	10.14			
BLD030B	0.35	0.00			
BLD520B	0.56	0.25			
BLDTR17	0.15	3.01			
CLN401C	4.37	5.08			
CLNAS43	7.06	4.00			
CLNAS98	2.46	2.28			
CLNCM12	1.83	3.18			
CLNDC19	23.78	3.17			
CLNRC01	1.61	7.21			
CLNRS53	2.52	4.25			

CLNSG27	5.93	3.42			
CLNTX01	1.95	4.68			
CVXKS52	14.0 6	12.7 8			
CVXNK23	4.85	5.22			
CVXNKS54	5.30	10.2 5			
CVXNKS55	30.9 0	21.0 5			
CVXNKS81	3.96	4.74			
ENDO10479	14.8 9	0.15			
ENDO28XA	18.6 1	3.72			
ENDO8XA	0.19	12.2 8			
KID106XD	0.48	0.24			
KID107XD	0.00	0.54			
KID109XD	0.33	1.11			
KID10XD	0.00	0.31			
KID22K	0.01	0.09			
LNG205L	0.20	2.52			
LNG315L	6.48	2.77			
LNG507L	9.69	6.93			
LNG528L	14.3 2	3.49			
LNG8837L	22.2 7	9.23			
LNGAC11	2.81	3.88			
LNGAC39	16.0 6	3.45			
LNGSQ80	5.54	1.67			
LNGSQ81	13.7 5	4.88			
LVR15XA	0.00	0.02			
LVR174L	0.00	0.01			
LVR187L	0.03	9.48			
OVRG021	2.78	0.06			
OVR10050	15.0 9				
OVR10400	24.0 6				
OVR1050	13.9 3				
OVR130X	17.4 3				
OVR7180	14.4 2				
OVRA1B	17.9 5				

OVR123O		0.60		
OVR18GA		0.00		
OVR206I		0.28		
OVR337O		0.00		
OVR40G		0.00		
OVR515O		0.00		
OVRC004		0.00		
OVRC177		0.00		
PAN71XL	5.62	2.65		
PAN82XP	0.49	4.10		
PAN92X	11.5 4	0.00		
PRO23B	8.43	6.39		
PRO65XB	4.54	8.83		
PRO675P	16.0 6	8.21		
PRO84XB	7.86	2.26		
PRO958P	8.48	11.0 5		
PRO263C			10.5 2	
PRO276P			4.63	
PRO767B			6.03	
PRO855P			4.85	
PRO10R				3.93
PRO20R				7.06
SKN287S	14.0 9	0.74		
SKN39A	0.25	0.00		
SKN669S	2.32	13.4 0		
SMINT171S	12.9 1	3.63		
SMINT20SM	16.4 7	7.21		
SMINTH89	10.4 9	2.25		
STO261S	8.68	4.96		
STO288S	2.53	2.07		
STO509L	3.07	4.69		
STO88S	2.32	3.78		
THRD143N	5.81	18.5 7		
THRD270T	6.26	6.37		
THRD56T	12.2 4	6.35		
TST39X	5.90	21.8 5		
TST647T	13.5 4	4.87		

TST663T	7.74	0.04			
UTR135XO	0.77	1.35			
UTR85XU	12.2 3	5.79			
BLOB1			0.00		
BLOB3			0.55		
BLOB5			43.6 0		
BLOB6			0.00		
BLOB11			0.00		
BLO982B			0.00		
ADR48AD			0.00		
CLN01CL			1.01		
CVX1ACV			6.51		
ES001ES			2.31		
HRT46HR			0.00		
HUMREF00H R	1.00				
KID55KD			0.18		
LVR89LV			0.02		
LNG90LN			11.6 7		
MAM01MA			3.96		
MSL84MU			0.00		
OVR3APV			0.02		
PAN04PA			3.52		
PLA59PL			3.27		
PRO09PR			2.29		
REC21RC			4.57		
SMINT59SM			1.50		
SPL7GSP			0.00		
STO09ST			1.79		
THYM99TM			1.05		
TRA16TR			16.1 2		
TST4GTS			0.41		
UTR57UT			1.26		

0.00= Negative or no expression

The sensitivity for Mam113 expression was calculated for the cancer samples versus normal samples. The sensitivity value indicates the percentage of cancer samples
5 that show levels of Mam113 at least 2 fold higher than the normal tissue or the corresponding normal adjacent form the same patient.

This specificity is an indication of the level of breast tissue specific expression of the transcript compared to all the other tissue types tested in our assay. Thus, these

experiments indicate Mam113 being useful as a breast cancer diagnostic marker and/or therapeutic target.

Sensitivity and specificity data is reported in the table below.

	CLN	LNG	MAM	OVR	PRO
Sensitivity, Up vs. NAT	11%	67%	80%	0%	20%
Sensitivity, Down vs. NAT	22%	11%	7%	0%	0%
Sensitivity, Up vs. NRM	67%	0%	47%	100%	80%
Sensitivity, Down vs. NRM	0%	33%	13%	0%	0%
Specificity	22.22%	31.75%	28.25%	31.41%	31.94%

5 Altogether, the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Mam113 a good marker for diagnosing, monitoring, staging, imaging and/or treating breast cancer.

10 Additionally, the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Mam113 a good marker for diagnosing, monitoring, staging, imaging and/or treating ovarian or prostate cancer.

Primers used for QPCR Expression Analysis of Mam113 are as follows:

(Mam113_forward): TGGTTGAGAAGACATGAAAATCCA (SEQ ID NO:233)

(Mam113_reverse): AATTCCACCCTGTCAACCTAAAAAA (SEQ ID NO:234)

(Mam113_probe): TGATTTGGTGTTCGAATTCAGGCAA (SEQ ID NO:235)

15

DEX0452_033.nt.1 (Mam128v2)

The relative expression level of Mam128v2 in various tissue samples is included below. Tissue samples include 70 pairs of matching samples, 7 non-matched cancer samples, and 34 normal samples, all from various tissues annotated in the table. A 20 matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. Of the normal samples 5 were blood samples which measured the expression levels in blood cells. Additionally, 2 prostatitis, and 3 Benign Prostatic Hyperplasia (BPH) samples are included. All the values are compared to breast cancer sample MAM355 (calibrator).

25 The table below contains the relative expression level values for the sample as compared to the calibrator. The table includes the Sample ID, and expression level values

for the following samples: Cancer (CAN), Normal Adjacent Tissue (NAT), Normal Tissue (NRM), Benign Prostatic Hyperplasia (BPH), and Prostatitis (PROST).

Sample ID	CAN	NAT	NRM	BPH	PROST
MAM355	1.00	0.00			
MAMS621	0.02	0.00			
MAMS516	0.00	0.00			
MAM522	0.13	0.00			
MAM76DN	0.20	0.00			
MAM976M	0.00	0.00			
MAM781M	0.00	0.00			
MAM19DN	0.00	0.00			
MAM517	0.00	0.00			
MAMS997	1.48	0.00			
MAM42DN	0.00	0.00			
MAM869M	6.48	0.00			
MAMS699	109.5 4	0.00			
MAMS570	27.16	0.00			
BLD030B	0.00	0.00			
BLD520B	0.00	0.00			
BLDTR17	0.00	0.00			
CLN401C	0.00	1.11			
CLNAS43	0.00	8.22			
CLNAS98	0.00	0.00			
CLNCM12	0.21	0.00			
CLNDC19	0.00	0.00			
CLNRC01	7.28	1.10			
CLNRS53	0.00	0.00			
CLNSG27	0.00	0.00			
CVXKS52	0.00	9.04			
CVXNK23	0.00	0.00			
CVXNKS54	0.00	0.00			
CVXNKS55	6.35	0.00			
CVXNKS81	0.00	0.00			
ENDO10479	59.21	0.00			
ENDO28XA	0.00	0.00			
ENDO8XA	0.00	0.75			
KID106XD	0.00	0.00			
KID107XD	0.00	0.89			
KID109XD	2.65	0.00			
KID10XD	0.00	1.15			
KID22K	0.00	3.50			
LNG205L	0.00	0.00			
LNG315L	0.00	0.00			
LNG507L	0.00	0.00			
LNG528L	3.04	0.00			

LNG8837L	0.00	0.00			
LNGAC11	0.00	0.00			
LNGSQ80	0.00	0.00			
LNGSQ81	0.00	0.00			
LVR174L	0.00	0.00			
LVR187L	0.00	0.00			
OVRG021	0.00	9.49			
OVR10050	0.00				
OVR10400	209.9 6				
OVR1050	0.00				
OVR130X	0.00				
OVR7180	10.08				
OVRA1B	3.11				
OVR1230		0.00			
OVR18GA		0.00			
OVR206I		0.00			
OVR3370		0.00			
OVR40G		14.3 7			
OVRC004		0.00			
OVRC177		17.7 7			
PAN71XL	3.25	0.00			
PAN92X	0.00	0.00			
PRO65XB	0.70	1.32			
PRO675P	0.00	0.00			
PRO84XB	0.56	0.00			
PRO958P	0.00	0.00			
PRO263C			0.0 0		
PRO767B			0.0 0		
PRO855P			0.0 0		
PRO10R				0.00	
PRO20R					1.99
SKN287S	0.00	0.00			
SKN669S	0.00	0.00			
SMINT171S	0.00	0.00			
SMINTH89	0.87	0.00			
STO261S	17.21	30.4 4			
STO288S	0.40	0.00			
STO88S	0.00	0.00			
THRD143N	2.84	0.36			
THRD270T	0.54	0.00			
THRD56T	337.9 9	0.00			

TST39X	0.00	0.00			
TST647T	0.00	3.13			
TST663T	7.58	0.65			
UTR135XO	7.62	16.7 5			
UTR85XU	0.00	11.1 0			
BLOB1		0.00			
BLOB3		0.00			
BLOB6		0.00			
BLOB11		0.00			
BLO982B		0.00			
ADR48AD		0.00			
BRN10BR		0.00			
CLN01CL		0.12			
ESO01ES		0.00			
HRT46HR		0.00			
HUMREF00H R	0.24				
KID55KD		0.04			
LVR89LV		0.00			
LNG90LN		0.23			
MAM01MA		0.00			
MSL84MU		0.00			
OVR3APV		0.03			
PAN04PA		0.31			
PLA59PL		9.05			
PRO09PR		0.43			
REC21RC		0.00			
SMINT59SM		1.39			
SPL7GSP		1.55			
STO09ST		0.00			
THYM99TM		6.79			
TRA16TR		17.4 3			
TST4GTS		0.00			
UTR57UT		0.00			

0.00= Negative or no expression

The sensitivity for Mam128v2 expression was calculated for the cancer samples versus normal samples. The sensitivity value indicates the percentage of cancer samples
5 that show levels of Mam128v2 at least 2 fold higher than the normal tissue or the corresponding normal adjacent form the same patient.

This specificity is an indication of the level of breast tissue specific expression of the transcript compared to all the other tissue types tested in our assay. Thus, these

experiments indicate Mam128v2 being useful as a breast cancer diagnostic marker and/or therapeutic target.

Sensitivity and specificity data is reported in the table below.

	CLN	LNG	MAM	OVR	PRO
Sensitivity, Up vs. NAT	22%	11%	53%	0%	20%
Sensitivity, Down vs. NAT	22%	0%	0%	0%	0%
Sensitivity, Up vs. NRM	11%	11%	53%	43%	0%
Sensitivity, Down vs. NRM	78%	78%	0%	0%	60%
Specificity	70.27%	68.11%	69.36%	71.12%	70.05%

5 Altogether, the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Mam128v2 a good marker for diagnosing, monitoring, staging, imaging and treating breast cancer.

Primers used for QPCR Expression Analysis of Mam128v2 are as follows:

(Mam128v2_forward): AGGGGGATTACAATGATGGACC (SEQ ID NO:236)

10 (Mam128v2_reverse): TTGCCAAGGTGCGAGCTT (SEQ ID NO:237)

(Mam128v2_probe): AGTGAGCGCTTAGATGCCAGCA (SEQ ID NO:238)

DEX0452_033.nt.2 (Mam128v3)

The relative expression level of Mam128v3 in various tissue samples is included 15 below. Tissue samples include 78 pairs of matching samples, 7 non matched cancer samples, and 35 normal samples, all from various tissues annotated in the table. A matching pair is formed by mRNA from the cancer sample for a particular tissue and mRNA from the normal adjacent sample for that same tissue from the same individual. Of the normal samples 5 were blood samples which measured the expression levels in blood 20 cells. Additionally, 2 prostatitis, and 4 Benign Prostatic Hyperplasia (BPH) samples are included. All the values are compared to breast cancer sample MAM355 (calibrator).

The table below contains the relative expression level values for the sample as compared to the calibrator. The table includes the Sample ID, and expression level values for the following samples: Cancer (CAN), Normal Adjacent Tissue (NAT), Normal Tissue 25 (NRM), Benign Prostatic Hyperplasia (BPH), and Prostatitis (PROST).

Sample ID	CAN	NAT	NRM	BPH	PROST

MAM355	1.00	0.0 9				
MAMB011X	0.05	0.7 0				
MAMS621	0.01	0.0 0				
MAMS516	0.01	0.0 0				
MAM522	1.68	0.2 4				
MAM76DN	0.50	0.1 6				
MAM976M	0.75	0.3 7				
MAM781M	0.94	0.2 7				
MAM19DN	0.18	0.6 1				
MAM517	0.88	0.0 0				
MAMS997	1.02	0.3 2				
MAM42DN	1.02	0.0 0				
MAM869M	0.71	0.3 5				
MAMS699	0.52	1.6 6				
MAMS570	0.49	0.8 8				
BLD030B	0.75	0.4 9				
BLD520B	1.95	0.5 7				
BLDTR17	0.37	0.3 8				
CLN401C	0.21	0.1 7				
CLNAS43	0.57	0.3 4				
CLNAS98	0.16	0.1 9				
CLNCM12	0.15	0.0 9				
CLNDC19	0.23	0.3 7				
CLNRC01	0.29	0.1 4				
CLNRS53	0.59	0.6 9				
CLNSG27	0.00	0.2 5				

CLNTX01	0.39	0.0 6				
CVXKS52	0.81	0.3 5				
CVXNK23	0.23	0.0 0				
CVXNKS54	0.62	0.2 2				
CVXNKS55	1.14	0.8 1				
CVXNKS81	0.31	0.0 0				
ENDO10479	0.48	0.6 7				
ENDO28XA	0.91	0.8 6				
ENDO8XA	0.55	0.5 4				
KID106XD	0.01	0.0 9				
KID107XD	0.42	0.1 5				
KID109XD	0.28	0.1 2				
KID10XD	0.12	0.1 1				
KID22K	0.11	0.0 4				
LNG205L	1.11	1.1 4				
LNG315L	0.37	1.2 0				
LNG507L	0.28	1.0 4				
LNG528L	1.75	1.4 4				
LNG8837L	0.70	0.9 1				
LNGAC11	0.57	0.9 6				
LNGAC39	4.55	0.8 8				
LNGSQ80	0.31	0.7 0				
LNGSQ81	0.57	0.3 0				
LVR15XA	0.14	0.2 6				
LVR174L	0.04	0.0 4				
LVR187L	0.00	0.4 6				

OVRG021	0.39	0.7 5			
OVR10050	0.87				
OVR10400	31.8 8				
OVR1050	0.40				
OVR130X	0.88				
OVR7180	0.79				
OVRA1B	0.89				
OVR1230		0.00			
OVR18GA		0.58			
OVR206I		0.81			
OVR3370		1.08			
OVR40G		1.32			
OVR5150		0.46			
OVRC004		14.7 7			
OVRC177		0.55			
PAN71XL	0.33	0.2 3			
PAN82XP	0.38	0.0 0			
PAN92X	2.32	2.4 9			
PRO23B	0.67	0.3 9			
PRO65XB	0.23	0.5 8			
PRO675P	0.29	0.3 0			
PRO84XB	0.30	0.9 3			
PRO958P	0.39	0.3 6			
PRO263C			0.0 7		
PRO276P			0.1 6		
PRO767B			1.0 1		
PRO855P			0.6 9		
PRO10R				0.60	
PRO20R				0.43	
SKN287S	0.20	0.6 9			
SKN39A	0.86	1.2 3			
SKN669S	0.67	0.6 1			

SMINT171S	0.51	0.1 4			
SMINT20SM	0.71	0.3 8			
SMINTH89	1.43	0.4 0			
STO261S	0.59	0.5 1			
STO288S	0.17	0.1 3			
STO88S	0.83	0.2 7			
THRD143N	0.29	0.8 4			
THRD270T	0.20	0.3 9			
THRD56T	0.38	0.1 9			
TST39X	0.20	0.0 0			
TST647T	0.94	0.5 6			
TST663T	0.47	0.5 3			
UTR135XO	1.25	1.4 0			
UTR85XU	1.39	2.9 3			
BLOB1		19.0 2			
BLOB3		1.62			
BLOB6		9.18			
BLOB11		3.51			
BLO982B		11.1 3			
ADR48AD		0.12			
BRN10BR		0.29			
CLN01CL		0.02			
ESO01ES		0.18			
HRT46HR		0.04			
HUMREF00H R	0.29				
KID55KD		0.01			
LVR89LV		0.03			
LNG90LN		0.11			
MAM01MA		0.01			
MSL84MU		0.03			
OVR3APV		0.02			
PAN04PA		0.00			
PLA59PL		0.25			
PRO09PR		0.86			

REC21RC		1.20		
SMINT59SM		0.21		
SPL7GSP		1.76		
STO09ST		0.06		
THYM99TM		0.31		
TRA16TR		0.59		
TST4GTS		0.36		
UTR57UT		1.65		

0.00= Negative or no expression

The sensitivity for Mam128v3 expression was calculated for the cancer samples versus normal samples. The sensitivity value indicates the percentage of cancer samples
5 that show levels of Mam128v3 at least 2 fold higher than the normal tissue or the corresponding normal adjacent form the same patient.

This specificity is an indication of the level of breast tissue specific expression of the transcript compared to all the other tissue types tested in our assay. Thus, these experiments indicate Mam128v3 being useful as a breast cancer diagnostic marker and/or
10 therapeutic target.

Sensitivity and specificity data is reported in the table below.

	CLN	LNG	MAM	OVR	PRO
Sensitivity, Up vs. NAT	22%	11%	67%	0%	0%
Sensitivity, Down vs. NAT	11%	33%	20%	0%	40%
Sensitivity, Up vs. NRM	89%	100%	87%	14%	0%
Sensitivity, Down vs. NRM	11%	0%	0%	0%	80%
Specificity	9.19%	14.05%	12.14%	15.51%	10.7%

Altogether, the tissue specificity, plus the mRNA differential expression in the samples tested are believed to make Mam128v3 a good marker for diagnosing,
15 monitoring, staging, imaging and treating breast cancer.

Primers used for QPCR Expression Analysis of Mam128v3 are as follows:

(Mam128v3_forward): ACAATAAATCAGTAAGCGTTCCAGAA (SEQ ID NO:239)

(Mam128v3_reverse): CAATCTACATTAAAACATACACGTGAACA (SEQ ID NO:240)

(Mam128v3_probe): CTTCTTCACCTCCTGAGGCCACTCA (SEQ ID NO:241)

Altogether, the high level of tissue specificity, plus the mRNA overexpression in matched samples tested are indicative of SEQ ID NO: 1-95 being a diagnostic marker and/or a therapeutic target for cancer.

Example 3: Protein Expression

5 The BSNA is amplified by polymerase chain reaction (PCR) and the amplified DNA fragment encoding the BSNA is subcloned in pET-21d for expression in *E. coli*. In addition to the BSNA coding sequence, codons for two amino acids, Met-Ala, flanking the NH₂-terminus of the coding sequence of BSNA, and six histidines, flanking the COOH-terminus of the coding sequence of BSNA, are incorporated to serve as initiating
10 Met/restriction site and purification tag, respectively.

An over-expressed protein band of the appropriate molecular weight may be observed on a Coomassie blue stained polyacrylamide gel. This protein band is confirmed by Western blot analysis using monoclonal antibody against 6X Histidine tag.

15 Large-scale purification of BSP is achieved using cell paste generated from 6-liter bacterial cultures, and purified using immobilized metal affinity chromatography (IMAC). Soluble fractions that are separated from total cell lysate were incubated with a nickel chelating resin. The column is packed and washed with five column volumes of wash buffer. BSP is eluted stepwise with various concentration imidazole buffers.

Example 4: Fusion Proteins

20 The human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector. For example, if pC4 (Accession No. 209646) is used, the human Fc portion can be ligated into the BamHI cloning site. Note
25 that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 2, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced. If the naturally occurring signal sequence is used to
30 produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. *See, e.g.*, WO 96/34891.

Example 5: Production of an Antibody from a Polypeptide

In general, such procedures involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to 5 culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100, µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; 10 however, it is preferable to employ the parent myeloma cell line (SP20), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.*, *Gastroenterology* 80: 225-232 (1981).

The hybridoma cells obtained through such a selection are then assayed to identify 15 clones which secrete antibodies capable of binding the polypeptide. Alternatively, additional antibodies capable of binding to the polypeptide can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific 20 antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein specific antibody and can be used to immunize an animal to 25 induce formation of further protein-specific antibodies.

Example 6: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

RNA is isolated from individual patients or from a family of individuals that have a phenotype of interest. cDNA is then generated from these RNA samples using protocols 30 known in the art. See, Sambrook (2001), *supra*. The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO: 1-95. Suggested PCR conditions consist of 35 cycles at 95°C for 30 seconds; 60-120 seconds at 52-58°C;

and 60-120 seconds at 70°C, using buffer solutions described in Sidransky *et al.*, *Science* 252(5006): 706-9 (1991). See also Sidransky *et al.*, *Science* 278(5340): 1054-9 (1997).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTTM Polymerase. (Epicentre Technologies). The 5 intron-exon borders of selected exons are also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations are then cloned and sequenced to validate the results of the direct sequencing. PCR products are cloned into T-tailed vectors as described in Holton *et al.*, *Nucleic Acids Res.*, 19: 1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected 10 individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements may also be determined. Genomic clones are nick-translated with digoxigenin deoxyuridine 5' triphosphate (Boehringer Manheim), and FISH is performed as described in Johnson *et al.*, *Methods Cell Biol.* 35: 73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 15 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C-and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and 20 variable excitation wavelength filters. Johnson (1991). Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

25 **Example 7: Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample**

Antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal 30 and are produced by the method described above. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced. The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial

dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbound polypeptide. Next, 50 µl of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three 5 times with deionized or distilled water to remove unbound conjugate. 75 µl of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution are added to each well and incubated 1 hour at room temperature.

The reaction is measured by a microtiter plate reader. A standard curve is prepared, using serial dilutions of a control sample, and polypeptide concentrations are 10 plotted on the X-axis (log scale) and fluorescence or absorbance on the Y-axis (linear scale). The concentration of the polypeptide in the sample is calculated using the standard curve.

Example 8: Formulating a Polypeptide

The secreted polypeptide composition will be formulated and dosed in a fashion 15 consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the secreted polypeptide alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

20 As a general proposition, the total pharmaceutically effective amount of secreted polypeptide administered parenterally per dose will be in the range of about 1, µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given 25 continuously, the secreted polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 mg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the 30 desired effect.

Pharmaceutical compositions containing the secreted protein of the invention are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally,

topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which

- 5 include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The secreted polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semipermeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-
10 release matrices include polylactides (U. S. Pat. No.3,773,919, EP 58,481, the contents of which are hereby incorporated by reference herein in their entirety), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U. et al., Biopolymers 22: 547-556 (1983)), poly (2-hydroxyethyl methacrylate) (R. Langer et al., J. Biomed. Mater. Res. 15: 167-277 (1981), and R. Langer, Chem. Tech. 12: 98-105 (1982)), ethylene vinyl acetate
15 (R. Langer et al.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release compositions also include liposomally entrapped polypeptides. Liposomes containing the secreted polypeptide are prepared by methods known per se: DE Epstein et al., Proc. Natl. Acad. Sci. USA 82: 3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA 77:
4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese
20 Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324, the contents of which are hereby incorporated by reference herein in their entirety. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal secreted polypeptide therapy.

25 For parenteral administration, in one embodiment, the secreted polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation.

30 For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides. Generally, the formulations are prepared by contacting the polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is

shaped into the desired formulation. Preferably, the carrier is a parenteral carrier, more preferably, a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as 5 liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic 10 acid; low molecular weight (less than about ten residues) polypeptides, e. g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such 15 as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The secreted polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers 20 will result in the formation of polypeptide salts.

Any polypeptide to be used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a 25 stopper pierceable by a hypodermic injection needle.

Polypeptides ordinarily will be stored in unit or multi-dose containers, for example, sealed ampules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1 % (w/v) aqueous polypeptide solution, and the resulting mixture 30 is lyophilized. The infusion solution is prepared by reconstituting the lyophilized polypeptide using bacteriostatic Water-for-Injection.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions

of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may 5 be employed in conjunction with other therapeutic compounds.

Example 9: Method of Treating Decreased Levels of the Polypeptide

It will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. 10 Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a pharmaceutical composition comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

For example, a patient with decreased levels of a polypeptide receives a daily dose 15 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided above.

Example 10: Method of Treating Increased Levels of the Polypeptide

Antisense or RNAi technology are used to inhibit production of a polypeptide of 20 the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer.

For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the 25 treatment was well tolerated. The formulation of the antisense polynucleotide is provided above.

Example 11: Method of Treatment Using Gene Therapy

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a 30 subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a

tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e. g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37°C for approximately one week.

5 At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks. pMV-7 (Kirschmeier, P. T. et al., DNA, 7: 219-25 (1988)), flanked by the long terminal repeats of the Moloney murine 10 sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5'and 3'end sequences respectively as set forth in 15 Example 3. Preferably, the 5'primer contains an EcoRI site and the 3'primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are 20 then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

The amphotropic pA317 or GP+aml2 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to 25 the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing 30 the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media.

If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

- 5 The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytadex 3 microcarrier beads.

Example 12: Method of Treatment Using Gene Therapy-In Vivo

- Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the
10 introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide.

- The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in
15 the art, see, for example, Tabata H. *et al.* *Cardiovasc. Res.* 35 (3): 470-479 (1997); Chao J *et al.* *Pharmacol. Res.* 35 (6): 517-522 (1997); Wolff J. A. *Neuromuscul. Disord.* 7 (5): 314-318 (1997), Schwartz B. *et al.* *Gene Ther.* 3 (5): 405-411 (1996); and Tsurumi Y. *et al.* *Circulation* 94 (12): 3281-3290 (1996); WO 90/11092, WO 98/11779; U. S. Patent No.
20 5,693,622; 5,705,151; 5,580,859, the contents of which are hereby incorporated by reference herein in their entirety.

- The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, breast, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

- 25 The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P. L. *et al.* *Ann. NY Acad. Sci.* 772: 126-139 (1995) and Abdallah B. *et al.* *Biol. Cell* 85 (1): 1-7 (1995)) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, 5 one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues 10 within the an animal, including of muscle, skin, brain, breast, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or 15 chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They 20 are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or 25 RNA will be in the range of from about 0.05 µg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. 30 The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also

be used, such as, inhalation of an aerosol formulation particularly for delivery to breasts or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

5 The dose response effects of injected polynucleotide in muscle *in vivo* is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

10 Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over one minute, 15 approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

15 After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual 20 quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.

25 The results of the above experimentation in mice can be used to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 13: Transgenic Animals

The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea 30 pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e. g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific

embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

Any technique known in the art may be used to introduce the transgene (I. e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., *Appl. Microbiol. Biotechnol.* 40: 691-698 (1994); Carver et al., *Biotechnology* 11: 1263-1270 (1993); Wright et al., *Biotechnology* 9: 830-834 (1991); and U. S. Pat. No. 4,873,191, the contents of which is hereby incorporated by reference herein in its entirety); retrovirus mediated gene transfer into germ lines (Van der Putten et al.,

- 10 *Proc. Natl. Acad. Sci., USA* 82: 6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., *Cell* 56: 313-321 (1989)); electroporation of cells or embryos (Lo, 1983, *Mol Cell. Biol.* 3: 1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e. g., Ulmer et al., *Science* 259: 1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells 15 and transferring the stem cells back into the blastocyst; and sperm mediated gene transfer (Lavitrano et al., *Cell* 57: 717-723 (1989). For a review of such techniques, see Gordon, "Transgenic Animals," *Intl. Rev. Cytol.* 115: 171-229 (1989).

- 20 Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., *Nature* 380: 64-66 (1996); Wilmut et al., *Nature* 385: 810813 (1997)).

- The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, I. e., mosaic animals or chimeric. The transgene may be integrated as a single transgene 25 or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., *Proc. Natl. Acad. Sci. USA* 89: 6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, 30 and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of

integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., *Science* 265: 103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

Example 14: Knock-Out Animals

Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E. g., see

Smithies et al., *Nature* 317: 230-234 (1985); Thomas & Capecchi, *Cell* 51: 503-512 (1987);

- 5 Thompson et al., *Cell* 5: 313-321 (1989)) Alternatively, RNAi technology may be used. For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that
- 10 express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used
- 15 to generate animal offspring with an inactive targeted gene (e. g., see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However, this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

- 20 In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e. g., knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts,
- 25 bone marrow cells, blood cells (e. g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral
- 30 vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc.

The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be
5 introduced into the patient systemically, e. g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e. g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U. S. Patent No. 5,399,349; and
10 Mulligan & Wilson, U. S. Patent No. 5,460,959, the contents of which are hereby incorporated by reference herein in their entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the
15 cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of
20 polypeptides of the present invention, studying conditions and/or disorders associated with aberrant expression, and in screening for compounds effective in ameliorating such conditions and/or disorders.

While preferred illustrative embodiments of the present invention are described, one skilled in the art will appreciate that the present invention can be practiced by other
25 than the described embodiments, which are presented for purposes of illustration only and not by way of limitation. The present invention is limited only by the claims that follow.

We claim:

1. An isolated nucleic acid molecule comprising:
 - (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-232;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b).
- 10
2. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a cDNA.
- 15 3. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is genomic DNA.
4. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is an RNA.
- 20
5. The nucleic acid molecule according to claim 1, wherein the nucleic acid molecule is a mammalian nucleic acid molecule.
6. The nucleic acid molecule according to claim 5, wherein the nucleic acid molecule is a human nucleic acid molecule.
- 25
7. A method for determining the presence of a breast specific nucleic acid (BSNA) in a sample, comprising the steps of:
 - (a) contacting the sample with the nucleic acid molecule of SEQ ID NO: 1-95 under conditions in which the nucleic acid molecule will selectively hybridize to a breast specific nucleic acid; and

(b) detecting hybridization of the nucleic acid molecule to a BSNA in the sample, wherein the detection of the hybridization indicates the presence of a BSNA in the sample.

5 8. A vector comprising the nucleic acid molecule of claim 1.

9. A host cell comprising the vector according to claim 8.

10. A method for producing a polypeptide encoded by the nucleic acid molecule
10 according to claim 1, comprising the steps of:

- (a) providing a host cell comprising the nucleic acid molecule operably linked to one or more expression control sequences, and
- (b) incubating the host cell under conditions in which the polypeptide is produced.

15

11. A polypeptide encoded by the nucleic acid molecule according to claim 1.

12. An isolated polypeptide selected from the group consisting of:

- (a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95.

25 13. An antibody or fragment thereof that specifically binds to:

- (a) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232 ; or
- (b) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95.

30 14. A method for determining the presence of a breast specific protein in a sample, comprising the steps of:

- (a) contacting the sample with a suitable reagent under conditions in which the reagent will selectively interact with the breast specific protein comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232; and
- 5 (b) detecting the interaction of the reagent with a breast specific protein in the sample, wherein the detection of binding indicates the presence of a breast specific protein in the sample.

15. A method for diagnosing or monitoring the presence and metastases of breast
10 cancer in a patient, comprising the steps of:

- (a) determining an amount of:
 - (i) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-232;
 - (ii) a nucleic acid molecule comprising a nucleic acid sequence of SEQ
15 ID NO: 1-95;
 - (iii) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (i) or (ii);
 - (iv) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (i) or (ii);
 - (v) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232 ; or
 - (vi) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95
20 and;
- (b) comparing the amount of the determined nucleic acid molecule or the polypeptide in the sample of the patient to the amount of the breast specific marker in a normal control; wherein a difference in the amount of the nucleic acid molecule or the polypeptide in the sample compared to the amount of the nucleic acid molecule or the polypeptide in the normal control is associated with the
25 presence of breast cancer.

16. A kit for detecting a risk of cancer or presence of cancer in a patient, said kit comprising a means for determining the presence of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-232;
 - 5 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b); or
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b); or
 - 10 (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95.

17. A method of treating a patient with breast cancer, comprising the step of administering a composition consisting of:
- (a) a nucleic acid molecule comprising a nucleic acid sequence that encodes an amino acid sequence of SEQ ID NO: 96-232;
 - 20 (b) a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
 - (c) a nucleic acid molecule that selectively hybridizes to the nucleic acid molecule of (a) or (b);
 - (d) a nucleic acid molecule having at least 95% sequence identity to the nucleic acid molecule of (a) or (b);
 - 25 (e) a polypeptide comprising an amino acid sequence with at least 95% sequence identity to of SEQ ID NO: 96-232 ; or
 - (f) a polypeptide comprising an amino acid sequence encoded by a nucleic acid molecule having at least 95% sequence identity to a nucleic acid molecule comprising a nucleic acid sequence of SEQ ID NO: 1-95;
- 30 to a patient in need thereof, wherein said administration induces an immune response against the breast cancer cell expressing the nucleic acid molecule or polypeptide.

18. A vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide of claim 12.

SEQUENCE LISTING

<110> diaDexus, Inc.
Macina, Roberto
Turner, Leah
Sun, Yongming
Chen, Huei-Mei
Rodriguez, Maria

<120> Compositions, Splice Variants and Methods Relating to Breast Specific Genes and Proteins

<130> DEX-0452

<150> US 60/431,123
<151> 2002-12-05

<160> 241

<170> PatentIn version 3.1

<210> 1
<211> 3163
<212> DNA
<213> Homo sapien

<400> 1
tccagtaagg tgtgccagc tttcttctgg gcacagacaa gaaagatgga aagtatacggt 60
taaagtcccc actctaaagt gctttacatt ttaaatgtgg accacaaaag tgcccacgag 120
ccaaaaagat tccaagaagc tgtgttaagca aatccatgat tgaatgttac aaacgtgtgt 180
ataaaagtgcgt gtgagatcag aaagccagga agtggctta aaaaaaaaaa actacaacca 240
aatcttcctt ctgggttctt caggcaattt ctctgcccac attcccattt ccctaagata 300
ttccataagg gccagtcacg gagaattcat acctgaaagg gaaactgttta tttgtgttgt 360
tgtcaaagat atgtggacta actttcagaa ctacccactg tgtttccttg gcaggttceg 420
gagtctcacc actgcatttt tcagagacgc catggcttc ttattaatgt ttgacctcac 480
cagtcAACAG agcttcttaa atgtcagaaa ctggatgagc caactgcaag caaatgctta 540
tttgtgaaaat ccagatatacg tattaattgg caacaaggca gacctaccag atcagaggga 600
agtcaatgaa cggcaagctc gggactggc tgacaaatat ggctgcaa at tgagtacact 660
ggaaatcaac aaatttgatg aagcctgtct gtctttcac cagtgagtg agtgcagcag 720
tttagaaagag aagcaatatt gtgcaactgg tgcaatggc agttaatcat agtgtataac 780
cttgcgttca tgaaacaggt tggtcattgt tctgcatttc tcttcattta aaaaggatac 840
acaattcttt cctcattgca tattacacca aacgtttgag ggaaaaatcc tcattcgtaa 900
aggatTTGG atttataatc taaaactcaa caataaagaa ataatattcc aagtctctgg 960
tttcctaaga tacataataa ctgtttataa agaaggctca agagctgata tttgccaaag 1020

tgatagaaga gttgtttttt cctctctact accaagcttt aagacattaa aagaagtcta	1080
gtgtatttga atattttaga gaaagcttta tcattttta agatgccaag atgctgccta	1140
cgtttcaaaa agttgtctaa gaattcacca tgagctatat tttcttcgg atcttgacc	1200
aagggtatgt cagcttattt ctggggagg tggtagctc ttatacatga aaatggatat	1260
aggctattct ctgggatgag tgtcatttca atgcttata aatccatgaa gctgctgtc	1320
tcataaaagta gaactgatac aaatttttgt tggatata gagaatttta taaatgtatt	1380
gccttagaat ttctgggtgg agacccaact acaatgacat tgtcatgaa gaactataaa	1440
gataattaga gttaaaagtt gttaaattt tgcccttaaa tacagcagaa cctggagaag	1500
gtcatacttc aaaggctcgat tttgagtcg aataaagaaa gacctagtaa cagatagttt	1560
tttttgttc atttcttct accaagtata ggtttatgcc ctcagaacta aactagtaaa	1620
aatatctgaa caaaaaacct ttcgttgg gcataaaaat gtgataact tagagacatt	1680
ttgtttattt catataaattc taattttcc ataaattaga tttatgatata tttcataaag	1740
cacttgatta gttttcaag gcgtaccatc acaaagatgc tttcctgcag agtttttgt	1800
atcaacagcc tatggttgag atgtttctc atttcctgta gagagagaat accactaaca	1860
aacaagcaaa aacttttagt cccaaatagt ggaactat tttcatctt tgaaaaaaa	1920
atatacaaag aagtcatctt ttcattaagt ggattccctg gttccttcc agctgggtgt	1980
ggaagtaatg gctaacatcc ttcagctgac tttgtctaca aggattatta gcaaattctg	2040
taggagcaag catgtctgac cttaacttaa tggatccctt attcaatcag tggcttctgt	2100
ctttatgtct gtggcatat caaaatggtt tctgttccata gaaaagtaat aacatatgct	2160
tatcttattt cttttccag gtgattttgt tttcaatgc tccttgcata aacacctagt	2220
gtttagaaaa gggaaagtggc cagaaagaac aacttgggac catgagtagg tcattaaata	2280
gcttagtgat ttatcctcat atagggotta taaaccctgt atgtgttat atgtgcttca	2340
cagagttcgt gtcaggctca aaggagatata gataagaaaa gtggttgtaa aattatgttc	2400
catttcataa atagacacta ttcacaaact aaaatctaataa aaaaaaccac agttgttaatt	2460
taaactgctt gatataaaaa gaggtatcat agcaggaaaa acacactaat tttcatacag	2520
tagaggtatt gaaaactgaa aatgggaagg caacttgaag tcattgtatt tgattaaaa	2580
tgtttaatac atctcattat tgacaaaata tgtcatctt tattttttc aaggaaacca	2640
atgaattcta ggttagtata tacaagttgg tcaaaatatt ccatgtacaa ataggcttc	2700
tgtgtccata gccttgcata agatactgat tgtatctgaa attatttttt aaaaaaataaa	2760
attatcctgc ttttagtgtgt taaaagtata cgatgttcta atataacact gaagtgcctc	2820
attgtatccc aacagtttac cttcaagtaa tattatctt attttttaggc taagcacgtt	2880

tgattatttt	gtctgtctcc	tatatagatc	tgtttgtct	agtgcstatga	atgtaactta	2940
aaactataaa	cttgaagttt	ttattctata	tgccccttaa	tagactgtgg	ttcctgacgc	3000
acactgttag	gtcattat	tgttgtacca	aagttctagt	ggcttcagaa	atcatagcat	3060
ccaatgattt	tttggtgtct	ggctatgaat	actatggttg	agaattgtat	tcagtgattg	3120
tttctgcaca	ctttcaa	aaaaatgaa	tttttatcaa	tta		3163

<210> 2
<211> 2305
<212> DNA
<213> Homo sapien

<400> 2						
taagctcgga	agcgactcta	gggcrggggg	agggtcgggt	gtcggcgagc	tccgcgtgcg	60
ggttccgagt	ggctgctggc	ggcctgggct	gccggggccg	acgcctgggt	ggctgctgcc	120
gccgcgcctg	ctgcgagatg	gcgatcttgg	gcgcggaaagg	gtgagggcgc	ccgcgcagg	180
aggaggtgcc	gctgccgtgg	ccgcggggct	gccgggagcc	gacagcttcg	cgcgggggtt	240
gtctcctcac	agactatgag	ctccttggaa	gagggaatcg	tgtcttactc	atctttgtat	300
ccccagtgtc	tagcagttcc	tgatacatag	tttagctga	attttggac	atggccactg	360
cttcaccaag	gtctgatact	agtaataacc	acagtggaaag	gttgcagtt	caggtaactg	420
tttctagtgc	caaactaaa	agaaaaaaga	actggttcgg	aacagcaata	tatacagaag	480
tagtttaga	tggagaaatt	acgaaaacag	caaaatccag	tagttcttct	aatccaaaat	540
gggatgaaca	gctaactgta	aatgttacgc	cacagactac	attggaaattt	caagtttgg	600
gccatcgcac	tttaaaagca	gatgcttat	tagaaaagc	aacgatagat	ttgaaacaag	660
ctctgttgat	acacaataga	aaattggaaa	gagtgaaaga	acaattaaaa	ctttccttgg	720
aaaacaagaa	tggcatagca	caaactggtg	aattgacagt	tgtgcttgat	ggattggtga	780
ttgagcaaga	aaatataaca	aactgcagct	catctccaac	catagaaata	caggaaaatg	840
gtgatgcctt	acatgaaaat	ggagagcctt	cagcaaggac	aactgccagg	ttggctgttg	900
aaggcacgaa	tggaatagat	aatcatgtac	ctacaagcac	tctagtccaa	aactcatgct	960
gctcgatgt	agttaatgga	gacaacacac	cttcatctcc	gtctcaggtt	gctgccagac	1020
ccaaaaatac	accagctcca	aaaccactcg	catctgagcc	tgccgatgac	actgttaatg	1080
gagaatcatc	ctcatttgca	ccaaactgata	atgcgtctgt	cacgggtact	ccagtagtgt	1140
ctgaagaaaa	tgccttgtct	ccaaattgca	ctagactac	tgttgaagat	cctccagttc	1200
aagaaataact	gacttcctca	gaaaacaatg	aatgtattcc	ttctaccagt	gcagaattgg	1260
aatctgaagc	tagaagtata	ttagagcctg	acacctctaa	ttctagaagt	agttctgctt	1320

ttgaagcagc caaatcaaga cagccagatg ggtgtatgga tcctgtacgg cagcagtctg	1380
ggaatgc当地 cacagaaaacc ttgccatcgag gtggggaaaca aagaaaagat cctcatggta	1440
gaaccttata tgtggatcat aataactcgaa ctaccacatg ggagagacca caacctttac	1500
ctccaggttg gaaaagaaga gttgatgatc gtagaagagt ttattatgtg gatcataaca	1560
ccagaacaac aacgtggcag cggcctacca tggaatctgt ccgaaatttt gaacagtggc	1620
aatctcagcg gaaccaattt caggagacta tgcaacagtt taaccaacqa tacctctatt	1680
cggcttcaat gtagctgca gaaaatgacc cttatggacc tttgccacca ggctggaaa	1740
aaagagtggta ttcaacagac agggtttact ttgtgaatca taacacaaaa acaacccagt	1800
gggaagatcc aagaactcaa ggcttacaga atgaagaaac cttggcaga aggctgcgac	1860
aatttagaat attctccgtg aaggtgctaa ggtcaccttg ctgcactcat tcaacccagc	1920
aacccacccc ctttccaaga ctccctcgca tgccggaaacc cactgacact tcaaacggtg	1980
gtccagcaaa ctgccttacc gAACGCCGGC tacaggtgaa gccagccaaa tacccaaaga	2040
tggggccccag cctaattggcc tacccacgca cgggaacgaa cacagcgtcc cccggccaac	2100
aatctgcgac ggaacccccc ccaacaaaga tggggcaaacc accccaagac agagaaggcc	2160
gccacagaaa ctttaccgca gagccagca ccaatcaggg cacgagaaaa gagccgaccc	2220
cacaacgtac caccacagt gcagacgcac aaccaactta gcaacgacaa caacacgaac	2280
actatacgca acaacacacaag caaca	2305

<210> 3
<211> 1900
<212> DNA
<213> Homo sapien

<400> 3	
tttttttat attttctaaa atttttatctt cttgttcatt ttgtttctaa gatattcact	60
cacatattaa aaataacaac gtctcaaaac atttgaagca actctttca tcccttttaa	120
aaatacccttg ctgtttcggg gttttttttt agccacaagg gagattttttt caatacaaaat	180
atttattttc ccaactcccc tggccatgggt tctgggacgt caccgcctt ttctggggcc	240
cgtttcatcc ttttctttta atccaagaag cgatgggttt gtgcgcctgt agtcccaagct	300
acctgggagg caggctgagg tgggaggatc ttttgggtcc aggattttga ggctgcagta	360
agccgttttc tcaccactgc actcgagact gggcgaaaga gcgagacact gtatcaaaaa	420
caaaaacaaaa caaaaacgaga aggcacgcg gctctgttac actccgttca gctctcgac	480
tctcagatgc aaacttccac acacacttccct cggctcgccct tttcccgccg gactagcata	540
tcaaggccttc cgggacacac cgtgcgtatga tatatacgta tataccccctc ttggcccttga	600

aggccggaag tcggtcttac agataaaagc gaaacaggaa gtcccgc(ccc tctatggaaa	660
gtaaatggta gtcggaaagg gtcaaaagag tcccggttt cgccgcgtga gttgctttt	720
gcggctgggg aggtctacgc ttcttagact tgagccagcg gggcgaccct gcagtggcag	780
gactcggcac cgccccc(cc accgcccgtt ggtggcctgc gtgacagttt cctccgtcg	840
acatcgaaag gaagccggac gtgggcgggc agagaggtcg gcttgctgat gggtccgggt	900
ggggcgcgcg tggactatgg gcccgggagg tcccttactg tccccgagcc gcgggttcct	960
cttgcacaa acggggtggc actccaatcg cctgcttggt gattgtggcc cccacacacc	1020
tgtttctaca gcgccttagct tcatacgact aggaatggca gccccatcta tgaaggaaag	1080
acaggtctgc tggggggccc gggatgagta ctggaagtgt ttagatgaga acttagagga	1140
tgcttctcaa tgcaagaagt taagaagctc tttcgaatca agttgtcccc aacagtggat	1200
aaaatatttt gataaaagaa gagactactt aaaattcaaa gaaaaatttg aagcaggaca	1260
atttgagcct tcagaaacaa ctgcaaaatc ctaggctgtt cataaagattt gaaagtattc	1320
tttctggaca ttgaaaagc tccactgact atggAACAGT aatagttga atcatagtga	1380
acatcaatac ttgttcccta tatacgacac ttgataatta agatgatcaa gaaccagaag	1440
atctgtgaag aaatgaaata aaatggtatt tagtaagaaa tctctatTTT aaaaaaaaaaa	1500
gtaaaaacctg ttataaacac aaaaaaaaaaa aaaaaaaaaaa aaaacaaaaa aaaaaaaaaaa	1560
aaacaaaaag aacggaacaa agacacaaaa aaaaacacaa cagaagaaag aggaaggggg	1620
ggggggcggg gggggcccgg gggggccggg gaccccccag ccccacccgg ggcgc(cccga	1680
tttagagagt ccatcacatc atccatcgct aagaaccacc agcaacacat gaacaacatg	1740
gattaaatac acatcacacc atgttatgtc ctctctaaga acaaccaaac gtatccgtag	1800
ctaacagtta gcaaactacg acatctatat gttcaatatt gattaatatt tggtaaaagt	1860
cagttgacaa tctctgtgat atcttgatac atttaacaa	1900

<210> 4
 <211> 1886
 <212> DNA
 <213> Homo sapien

<400> 4	
ttttttttgg aaccatgtgc gccttattttt gctgagccac tacttgagag ggatgaagca	60
gaaggagtggtt gttggccgatc tgccggaccg gcattgtttt acggggcttgtt aggtgtatgg	120
gaactcgccc aggttagtggc caatcatctc gggcttgacc tccacctgggt tgaaggtctt	180
ggcggtgtgg acgcccacca tgctgcccac cacctcgggc aggatgatca cgtccctcag	240
gtgcgtcttc accacttcgg gcttctccat gggcgccgccc tcccttgg ctttgccgcag	300

gcgcttcagc	agggagtgct	gcttccgccg	caggccccgg	ttcagccgcc	gcgctggcgc	360
gcactgtaca	gctgcatca	ctgctcgtag	gacatgtcca	gcagctggtc	gaggtccacg	420
ccgcgttca	tcgcagttagg	aatggcagcc	ccatctatga	aggaaagaca	ggtctgctgg	480
ggggcccccgg	atgagtactg	gaagtgttta	gatgagaact	tagaggatgc	ttctcaatgc	540
aagaagttaa	gaagctcttt	cgaatcaagt	tgtccccaac	agtggataaa	atattttgat	600
aaaagaagag	actacttaaa	attcaaagaa	aaatttgaag	caggacaatt	tgaggcattca	660
gaaacaactg	caaaatccta	ggctgtcat	aaagattgaa	agtattcttt	ctggacattg	720
aaaaagctcc	actgactatg	gaacagtaat	agtttgaatc	atagtgaaca	tcaatacttg	780
ttccctatat	acgacacttg	ataattaaga	tgtcaagaa	ccagaagatc	tgtgaagaaa	840
tcaaataaaaa	tgttatTTAG	taagaaatct	ctatTTAAG	aaaaaaaaagta	aaacctgtta	900
taaacatcg	cactTTTGT	ttgtTTTGT	tttgtTTTA	attagaggat	gggttagtagg	960
cagatgataa	aatttataat	atacatagaa	gtgaaataaa	tgggagttag	catttataa	1020
caggcaagag	ctattacaac	aacccaagtg	agaaatgatg	agggtttgtg	gaaggTTTAT	1080
aaggaagaag	ggtgaactta	aaatatacaa	gtaaaataat	aaaagccatc	tataaaaaag	1140
cccatagcta	atatcaacac	ttaatgttgg	gacaggaact	ggatgtctag	ctagtcagt	1200
gagacaaaaa	agaaaaagca	tacacactgg	gaaggaagaa	agaaaaactag	ctctactcac	1260
atataataaa	tactatctta	tagaatgtac	caatggatgc	acaaaaagag	ctcctagaac	1320
tataagtcaa	tcatagaaag	gttgcaggaa	acaaggtcaa	tatacacaag	gaaaattata	1380
ttcctatata	ttagcaataa	acaactggaa	tttAAAACtt	aaaaatacca	tttgtgaata	1440
gcaccaaaaa	aaattaaagg	aatacttagg	tataaatcta	atatatggag	gcctctatgc	1500
tgagaactag	aaaacacttg	gaagcagact	acatcagatt	aaatggagag	gtatacagtt	1560
ggccctctgt	gggttctgca	tccatggatt	caaccccgaa	gagaaaattt	ttgggaaaag	1620
gaaaaacgag	taaaaataat	aaaaattaa	aaatccagta	taacacctat	ttacattgta	1680
ttaggtattg	taagtcattg	agatgattta	aagtataggc	atacctcaaa	gatactgcag	1740
gtttggttac	agaccactgc	attaaagtga	atatcacaat	agagtgggtt	acacaaatgt	1800
tttggTTTC	cagtacacat	agaagttatg	tttatactgt	tgtctagtaa	gtgtgcaata	1860
gcattatgtc	tgcTcagtat	atatgc				1886

```

<210> 5
<211> 1935
<212> DNA
<213> Homo sapien

```

<400> 5	
agatccaaga tggcattat attcattgt a ttttacaaa ttcttacatt ttagttatc	60
ttcagcaaaa aatccagatg gatgttttt tcagaaagtg ttgaatgggt ttacaaagtt	120
ttttttaag gaacaatatt gcaaattact aaaattgtat ttttataggc tggttgctct	180
tttgtggata ttgtgcctgt caggattctt gaagttttt ttttatagtg agataatgga	240
gttggctta gccgctgcag gagccctct tttctgtgga ttcatcatct atgacacaca	300
ctcaactgatg cataaaactgt cacctgaaga gtacgttata gctgccatca gcctctactt	360
ggatatcatc aatctattcc tgcacctgtt acggttctg gaagcagtta ataaaaagta	420
ataaaaagta tctcagctca actgaagaac aaaaaaaaaa attaacgag aaaaaaggat	480
taaagtaatt ggaagcagta tatagaaact gttcattaa gtaataaagt ttgaaacaat	540
gattaaatac ttttacaatc ttttttta tcatatgtaa ttttgagagc tttaaaatct	600
tactattctt tatgataacct catttctaaa tccttgattt aggatctcag ttaagagcta	660
tcaaaattct ataaaaatg ctttctggc tggcacagt ggctcacgccc tgtaatcccc	720
ccactttggg agaccgagge aggtggatca cgaggtcaag aggttgagac catcctggcc	780
aacatggta aaccccgctc ctactaaaa tacaaaaatt agctggatgt ggtggcacac	840
acctgttagtc ccagctagtc aagaggctga ggccagagaa tcgcttgaac ctgggaggtg	900
gaggttgcatt tggccaaga tcacgccact gcattccagc ctggtgacag agcgagactc	960
agtctcaaaa aaaaaaaaaa aatttttctt cctaaattag ccacgcatacg cggttcgttt	1020
gcaattcaaa aataattttt tggtagata agaatatcag tttaccgtt tctagtgatt	1080
ttatctaaat tttccctgaa ttattaagta atattgattt ggctttgatt ctgaagtagt	1140
agagtcttta ccattataaa ctgtaaatct cttttgctt aaaaggaaaa aaatgtaaaa	1200
gataaattcc acagagaatt attcagtatt acattaaaat gtaatgact ttttattttt	1260
aattgtacta acattaaaag ttggcctgaa agtcagatattatgacaaaa tttgacattt	1320
attgtttta aagtagatattt ttcatttgaa attatagaat gtaatgtgg ttagaggaca	1380
ccaaagatac tgggtcatca gccattaagt atatctattt caaaattaaa atatttggga	1440
agtattgtct tatggttca tttgtgtgg tccacacagc atgttaggtc agtgtaccag	1500
taaccaatga aattttgtca aattccctca ctgtactagt ttgttaggtc gccataacaa	1560
agttctacag cttgggtggc ttcaacaaga aattttttt cccacagtcc tggaggctaa	1620
aagtccaaaga tcaagggtttt agcagggtt gttcccttgg aggcctttctt ctttgattt	1680
tagatggcca ttttccctt gtgttttaa atggccttcc ctctgtactt gtctgtgccc	1740
aaatttcttc ttcttatgag gacaccagtc atactggatt agggcccaca ctgaggaccc	1800

cattttcct taattatctc tttcaaaacc tatctccaaa tacagtaca ttctgaagtg 1860
 ctgggattag gatttcttca tgtgaatttt gggggacta caactcagcc cataacaccc 1920
 cctaagtatt tccca 1935

<210> 6
 <211> 2028
 <212> DNA
 <213> Homo sapien

<400> 6
 cccgtttcca tggctgcgag aactgaegct ccccaaccgt cccgcaactg tcctgtccca 60
 gactttggca ccgtcggggt ccgtcgccc cgaatgtgac agcatcccc ccccggtgc 120
 tgcccaggat ccggccggacc ccggcctcga tatgggagac ctggaaactgc tgctgcccgg 180
 ggaagctgaa gtgctggtgc ggggtctgcg cagttccc ctagcgcaga tgggctccga 240
 agggtggaac cagcagcatg agaacctgga gaagctgaac atgcaagcca tcctcgatgc 300
 cacagtcagc cagggcgagc ccattcagga gctgctggc acccatggga aggtcccaac 360
 actggtgag gagctgatcg cagtgagat gtggaaagcag aaggtgttcc ctgtgttctg 420
 cagggtgag gacttcaagc cccagaacac cttcccccatt tacatggtgg tgcaccacga 480
 ggcctccatc atcaacacct tggagacagt gttttccac aaggaggtgt gtgagtcagc 540
 agaagacact gtcttggact tggtagacta ttgccaccgc aaactgaccc tgctggtggc 600
 ccagagtggc tgggtggcc cccctgaggg ggagggatcc caggacagca accccatgca 660
 ggagctgcag aagcaggcag agctgatgga atttgagatt gcactgaagg ccctctcagt 720
 actacgctac atcacagact gtgtggacag cctctcttc acgacactttga gccgtatgct 780
 tagcacacac aacctgcctt gcctccttgtt ggaactgctg gagcatagtc cctggagccg 840
 gcgaaaaagga ggcaagctgc agcagttcga gggcagccgt tggcatactg tggcccccctc 900
 agagcagcaa aagctgagca agttggacgg gcaagtgtgg atcgccctgt acaacctgct 960
 gctaagccct gaggttcagg cgcgctactg cctcacaagt tttgccaagg gacggctact 1020
 caaggtcaga ctccctccgc accagcccc acagccccag taccgccttc cccatcctac 1080
 cccgactgcg tccctgtgt ttatcttgc ccaccacct caacccactg gctctttca 1140
 gtccttggc ctcaggtgac acaccagcta gtgggacatg ggcccccaca ggcattctca 1200
 gcccccaaccctt gccccttgc cccctggcca gcacctgcat cacactggcc 1260
 tccactggac acccttgcag ctccggccctt ccctcacaga cacactgctg gaccagctgc 1320
 ccaacctggc ccacttgcag agtttcttgg cccatctgac ccttaactgaa acccagccctc 1380
 ctaagaagga cctgggttttgaacagatcc cagaaatctg ggagcggctg gagcgagaaa 1440

acagaggcaa gtggcaggca attgccaagc accagctcca gcatgtgttc agccccc tag	1500
agcaggacct gcggctgcag gcgcgaaggt gggctgagac ctacaggctg gatgtgctag	1560
aggcagtggc tccagagcgg ccccgctgtg cttactgcag tgca gaggct tctaagcgct	1620
gctcacgatg ccagaatgag tggtattgct gcagggagtg ccaagtcaag cactggaaa	1680
agcatggaaa gacttgtgtc ctggcagccc agggtgacag agccaaatga gggctgcagt	1740
tgctgagggc cgaccaccca tgccaaggaa atccacccag aatgcacccc tgaacctcaa	1800
gatcacggtc cagcctctgc cggagccccaa gtctccgcag tggagagcag agcgggcggt	1860
aaagctgctg accgatctcc ctccctctca ccccaagtga aggctcgaga cttcctgcucc	1920
cacccagtgg gtagggcaag tgtgttgc tt cagcaaacgg gaccaggagg gccaggccc	1980
gatgtgggaa ccctcttcct ctagcacagt aaagctggcc tccagaaa	2028

<210> 7
<211> 3186
<212> DNA
<213> Homo sapien

<400> 7	
agatcaaaga aggaaagaag agaagggaa gaaaagaagg ggaagaagat caaaacccac	60
catgccccag gctcagcagg gagctgctgg atgagaaaga gcctgaagtc ttgcaggact	120
cactggatag atgttattca actccttcag gttgtgttga actgtgtgac tcataccgc	180
cctacagaag tgccctttat gtattggagc aacagcatgt tggcttgct gttgacatgg	240
atgaaattga aaagtaccaa gaagtggaaag aagaccaaga cccatcatgc cccaggctca	300
gcagggagct gctggatgag aaagagctg aagtcttgc ggactcaactg gatagatgtt	360
attcgactcc ttca gtttat cttaactgc ctgacttagg ccagccctac agcagtgctg	420
tttactcatt ggaggaacag taccttgct tggcttgc cgtggacaga attaaaaagg	480
accaagaaga ggaagaagac caaggccccac catgccccag gctcagcagg gagctgctgg	540
aggttagtaga gcctgaagtc ttgcaggact cactggatag atgttattca actccttcca	600
gttgttgc acagcctgac tcctgccagc cctatggaaat ttccctttat gcattggagg	660
aaaaacatgt tggctttct cttaactgc gagaaattga aaagaaggaa aagggaaaga	720
aaagaaggaa aagaagatca aagaaggaaa gaaagaaggaa gaagaaaaga agggaaagaa	780
gatcaaaacc caccatgccc caggctcagc agggagctgc tggatgagaa agggcctgaa	840
gtcttgcagg actcaactgga tagatgttac tcaactcctt cagttgtct tgaactgact	900
gactcatgcc agccctacag aagtgcctt tatgtattgg agcaacagcg tggatggcttgc	960
gctgttgcaca tggatgaaat tgaaaagtac caagaagtgg aagaagacca agaccatca	1020

10

tgccccaggc tcagcaggga gctgctggat gagaaaagagc ctgaagtctt gcaggactca 1080
 ctggatagat gtattcgac tccttcaggt tatcttgaac tgcctgactt aggccagccc 1140
 tacagcagtg ctgtttactc attggaggaa cagtagcttg gcttggctct tgacgtggac 1200
 aaaattgaaa agaaggggaa ggggaaaaaa agaaggggaa gaagatcaa gaaggaaaga 1260
 agaaggggaa gtaaagaagg ggaagaagat caaaaacccac catgccccag gctcagcggt 1320
 gtgctgatgg aagtggaga gcctgaagtc ttacaggact cactggatacg atgttattcg 1380
 actccgtcaa tgtactttga actacctgac tcattccagc actacagaag tgtgtttac 1440
 tcatttgagg aacagcacat cagttcgcc cttgacgtgg acaataggtt tcatttttg 1500
 atgggaacaa gtctccacct ggtcttccag atgggagtca tattcccaca gtaaggcagcc 1560
 cttactaagc cgagagatgt cattcctgca ggcaggacat ataggcacgt gaagatttg 1620
 atgaaactat agttccattt ggaagcccag acataggatg ggtcagtggg catggctcta 1680
 ttcctattct cagaccatgc cagttggcaac ctgtgctcag tctgaagaca atggacccaa 1740
 gtttaggtgtg acacgttcac ataactgtgc agcacatgcc gggagtgtac agtcagacat 1800
 ttttaatttg accacgtatc tctgggttagc tacaaagttc ctcagggatt tcatttgca 1860
 ggcatgtctc tgagcttcta tacctgctca aggtcagtgt catcttggt ttagctcat 1920
 ccaaagggtgt tacccctgggt tcaatgaacc taacctcatt ctttgtatct tcagtgttg 1980
 attgttttag ctgatccatc tttaacacag gaggatcct tggctgagga ttgtatttca 2040
 gaaccaccaa ctgctttga caattgttaa cccgctaggc tccttgggt agagaagcca 2100
 cagtccttca gcctccaatt ggtgttagta cttaggaaga ccacagctag atggacaaac 2160
 agcattggga ggcccttagcc ctgctccctc cgattccatc ctgttagagaa caggagtcag 2220
 gagccgctgg caggagacag catgtcaccc aggactctgc cggcgcagaa tatgaacaac 2280
 gccatgttct tgcaaaaaac gcttagcctg agttcatag gaggtaatca ccagacaact 2340
 gcagaatgtg gaacactgag caggacaact ggcctgtctc cttcacatag tccatatacac 2400
 cacaatcac acaacaaaaa ggagaagaga tattttgggt tcaaaaaaaag taaaaagata 2460
 atatagctgc atttcttttag ttatttgaa ccccaaatat ttcctcatct ttttgtgtt 2520
 gtcattgatg gtggtgacat ggacttggtt atagaggaca ggtcagctgt ctggctcagt 2580
 gatctacatt ctgaagttgt ctgaaaatgt cttcatgatt aaattcagcc taaacgtttt 2640
 gcccggaaaca ctgcagagac aatgctgtga gtttccaacc ttagccatc tgcggcaga 2700
 gaaggtctag tttgtccatc agcattatca tgatatcagg actggttact tggtaagga 2760
 ggggtctagg agatctgtcc ctttttagaga caccttactt ataatgaagt atttggagg 2820
 gtggtttca aaagtagaaa tgtcctgtat tccgatgatc atcctgtaaa cattttatca 2880

tttattaatc atccctgcct gtgtcttata ttatattcat atctctacgc tggaaaacttt	2940
ctgcctctat gtttactgtg cctttgttt tgctagtgtg tgggttggaa aaaaaaaaaaca	3000
ttctctgcct gagtttaat ttttgccaa agttttta atctatacaa taaaaagctt	3060
ttgcctatca aaaaaaaaaaag gggggggtaa aataccgagg gccaaattgg tccctttgt	3120
aaagggcctc aggagggtaa aagcagaggg gggtaacgga gggaaagcgca ggatgagaac	3180
tggggaa	3186

<210> 8
<211> 790
<212> DNA
<213> Homo sapien

<400> 8	
gctttgtctg tgtgatctgt gtgtgtatgt tgctttggga atcctgcccc gtgcagttta	60
ggaggagctc caggagkctg ctgkctggct cagagtctgt ccccggttat ccactagccc	120
agagcagttc tccctatagc ccagtaagaa attacacctt caccccccac actggcaccc	180
acgctctccc agaaagttag aagggaaactc acaggtgact tcaccccatg gtggggagaa	240
cagcctgtgc tggggtaaag gcagaaggag gatgagcccc gaggctctg gagagtctga	300
gcctgggtga ggaaggggag gaggtggtcc ctgatctcag ggccgggaga gccaatgagg	360
agacggagcc atagcacgca gctctcagct gggggatcct ggtccccctca ccatctcctc	420
tccccccagct actccgtgaa gtctaggac aggaagatgg ttggcgacgt gaccggggcc	480
caggccatag cctccaccgc caagtgcctg aacatctggg ccctgattct gggcatcctc	540
atgaccattt gattcatctt gttactggta ttccggctctg tgacagtctm ccattttatg	600
ttwcagataa tacagggaaaa acggggttac tagtagccgc ccatacgctg caacctttgc	660
actccactgt gcaatgctgg ccctgcacgc skggctgttgc cccctgcccc cttggctctg	720
cccttarata cagcagttta taccacacaca cctgttytaca gtgtcatcata ataaagcgca	780
cgtgcttgg	790

<210> 9
<211> 1233
<212> DNA
<213> Homo sapien

<400> 9	
tgcacgactc cggctggca ggattccgga caacgcctgg ttccctttgg gtccttccgg	60
cgtcgccgga gtgaattgtat ccggggatgg aagagggctg caaggtggga agtgaagtca	120
gtgcctcagt tgctgatcag tttttttt gtgtccaaatt cttttatcac caaaaaagag	180

12

aagaaaatatt	gcagtgaatg	aagattcctc	tgcattttag	cactgctttt	tcaactgttag	240
ttggcctttt	aatgaggatg	acaatggaag	agatgaagaa	tgaagctgag	accacatcca	300
tggtttctat	gccccccttat	gcagtcatgt	atccctgttt	taatgagcta	gaacgagtaa	360
atctgtctgc	agccccagaca	ctgagagccg	ctttcatcaa	ggctgaaaaa	gaaaatccag	420
gtctcacaca	agacatcatt	atgaaaattt	tagagaaaaa	aagcgtggaa	gttaacttca	480
cggagtcct	tcttcgtatg	gcagctgatg	atgtagaaga	gtatatgatt	gaacgaccag	540
agccagaatt	ccaagaccta	aacgaaaagg	cacgagcact	taaacaattt	ctcagtaaga	600
tcccagatga	gatcaatgac	agagtgaggt	ttctgcagac	aatcaagcac	ttgaacacca	660
aaagaaaagaa	tttgtaaagt	actccaaaag	tttcagtgtat	actctgaaaa	cgtattttaa	720
agatggcaag	gcaataaaatg	tgttcgtaaag	tgccaaaccga	ctaattcatc	aaaccaactt	780
aataacttcag	accttcaaaa	ctgtggcctg	aaagttgtat	atgttaagag	atgtacttct	840
cagtggcagt	attgaactgc	ctttatctgt	aaatttaaa	gtttgactgt	ataaattatc	900
agtccctcct	gaagggatct	aatccaggat	gttgaatggg	attattgcca	tcttacacca	960
tatTTTgtA	aaatgttagct	taatcataat	ctcacactga	agatTTGCA	tcactTTGc	1020
tattatcatt	cttttaagaa	ttataagcca	aaagaattt	cgccttaatg	tgtcattata	1080
taacattcct	taaaagaatt	gtaaatattt	gtgtttgttt	ctgacatttt	aacttgaaag	1140
cgatatgctg	caagataatg	tatTTAACAA	tatTTGGTGG	caaATATTCA	ataaaatagtt	1200
tacatctgtt	aaacatttct	ttacttgaaa	aaa			1233

<210> 10
<211> 596
<212> DNA
<213> Homo sapien

<400> 10	ggaagagttc	cccttgcttt	aggagtgcag	actctgcctc	aaacttgtga	tgaacccaaa	60
	gccccacacca	gcttccaaat	ctccctaagt	gtcagttaca	cagggtcgag	cgcccgcccc	120
	ggcaggtacg	aactgttcaa	gagctcccc	cactccctgt	tcccagaaaa	aatggctcc	180
	agctgtctcg	atgcacacac	tggtatatcc	catgaagacc	tcatccagg	ggggggaccc	240
	cccatttcac	tgcagattca	cgactcccc	gcattggcca	gtgcttctcc	acccttaagt	300
	cctgtgcctc	ccctctatgt	tgtagaaaga	gccaaatcac	agtccctgtgt	gactggggac	360
	agtcaacttc	cctgcctgag	catcagttt	ttctattaaa	tggggggcag	aaatgcatgt	420
	ggagcatttc	cttgtaaaaa	cctgagggtg	ggctgggcac	ggtggctcat	gcctataatc	480
	ccagcacttt	gggaggctga	ggcggggagga	ttacttaagc	ctagaagttt	gagagttga	540

gaccagcctg ggcaacataa tgagacctcg tctctccaaa aaaaaaaaaa aaaaaaa	596
<210> 11	
<211> 1674	
<212> DNA	
<213> Homo sapien	
<400> 11	
ctggctggcg tccctctcg cagggtgctc tttgccccat ggggtgggat ccagagctgc	60
agacaggccc ccaggcttgg ccaatgaaca gaccaggttc ggggagggtg ttggaaaaga	120
gtggatgggg tggttccct taccttgcag cccccaggcc ctccccccct ccctcccagg	180
tggtcgggac tcttgatctt cgctcggtt actgtctgtt cggctgtctt ccccgccctt	240
ccccaggcac ctgcattcctc ccttggcacc tgctgccagg ctaggaaggg caaaaacaat	300
cccagttggc gtagtcaggg agtctccgcc ctccctccag gtttcctcct cccaggcgcc	360
tcccttggac ccgccccat ctgccccaga taattttagt ttcccttggc ctggaatctg	420
gacacacagg gctccccccc gcctctgact tctctgtccg aagtccggac accctcctac	480
cacctgtaga gaagcgggag tggatctgaa ataaaatcca ggaatctggg ggttcctaga	540
cgagccaga cttcggAACG ggtgtcctgc tactcctgct ggggctcctc caggacaagg	600
gcacacaact ggttccgtta agccctctc tcgctcagac gccatggagc tggatctgtc	660
tccacccat cttagcagct ctccggaaaga cctttggcca gcccctggga cccctcctgg	720
gactccccgg cccctctgata cccctctgcc tgaggaggt aagaggtccc agcctcctt	780
catcccaacc accggcagga aacttcgaga ggaggagagg cgtgccacct ccctccctc	840
tatcccaac ccctccctg agctctgctg tccctccctca cagagccaa ttctcgaaaa	900
ccctccctgt gcaagggggc tgctcccccg cgatgccagc cgccccatg tagtaaaggt	960
gtacagttag gatggggcct gcaggtctgt ggaggtggca gcaggtgcca cagctcgcca	1020
cgtgtgtgaa atgctgggtgc agcgagctca cgccttgagc gacgagacct gggggcttgt	1080
ggagtgccac ccccacccatg cactggagcg gggtttggag gaccacgagt ccgtgggtgg	1140
agtgcaggct gcctggcccg tggggggaga tagccgcttc gtcttccggaa aaaacttcgc	1200
caagtacgaa ctgttcaaga gctccccaca ctccctgttc ccagaaaaaa tggtctccag	1260
ctgtctcgat gcacacactg gatatccca tgaagacctc atccaggtgg ggggaccccc	1320
catttcactg cagattcactg actccccaggc attggccagt gtttctccac ccttaagtcc	1380
tgtgcctccc ctctatgttg tagaaagagc caaatcacag tccctgtgtga ctggggacag	1440
tcactttccc tgccctgagca tcagtttctt ctattaaatg gggggagagaa atgcattgtgg	1500
agcatttccct tgtaaaaaacc tgagggtggg ctgggcacgg tggctcatgc ctataatccc	1560

14

agcactttgg	gaggctgagg	cgggaggatt	acttaagcct	agaagttga	gagtttggaa	1620
ccagcctggg	caacataatg	agacctcgtc	tctccaaaaa	aaaaaaaaaa	aaaa	1674
<210>	12					
<211>	2297					
<212>	DNA					
<213>	Homo sapien					
<400>	12					
agagttgggtt	tgttagtaact	ggcactcagg	aacatgaggg	aaaaaaatta	catattgtga	60
aatgggttag	aagacatgaa	aatccacttg	attttggtgt	ttccgaattt	caggcaaaga	120
actgtttttt	aggttgacag	ggtggaaattc	agataacttct	atgcattaaac	tgtataatca	180
aaaggaaattt	gcttgggata	ggataaaagaa	ctgtggtctc	tttctttaaa	atgtgttagat	240
ggaacagtga	ctatgttttt	agtgttagca	cgtgcatgtc	agctgttaca	aatatgtctc	300
aaagaatctc	tctttgcata	tctaggcctg	tctcctccct	cctacacatt	tccagctcct	360
gctgcagttt	ttcctacaga	agctgccatt	taccagccct	ctgtgatttt	aatccacgaa	420
gcactgcagc	cctccacagc	gtactaccca	gcagggactc	agctcttcat	gaactacaca	480
gcgtactatc	ccagcccccc	aggttcgcct	aatagtcttg	gctacttccc	tacagctgct	540
aatcttagcg	gtgtccctcc	acagcctggc	acgggtggca	aatgcaggg	cctggcctac	600
aatactggag	ttaaggaaat	tcttaacttc	ttccaagggtt	accagtatgc	aaccgaggat	660
ggacttatac	acacaaatga	ccagggcagg	actctaccca	aagaatgggt	ttgtattttaa	720
ggggccccagc	agttagaaca	tcctcagaaa	agaagtgttt	gaaagatgta	tggtgatctt	780
gaaacctcca	gacacaagaa	aacttcttagc	aaattcaggg	gaagtttgc	tacactcagg	840
ctgcagtttt	ttcagcaaac	ttgattggac	aaacgggcct	gtgccttatac	ttttgggtgga	900
gtgaaaaaat	ttgagctagt	gaagccaaat	cgtaacttac	agcaagcagc	atgcagcata	960
cctggctctt	tgtgatttgc	aaataggcat	ttaaaatgtg	aatttggaaat	cagatgtctc	1020
cattacttcc	agttaaagtg	gcatcatagg	tgtttcctaa	gttttaagtc	ttggataaaa	1080
actccaccag	tgtctaccat	ctccaccatg	aactctgtta	aggaagcttc	atttttgtat	1140
attcccgctc	ttttcttcttc	atttccctgt	cttctgcata	atcatgcctt	cttgctaagt	1200
aattcaagca	taagatcttg	gaataataaa	atcacaatct	taggagaaag	aataaaatttg	1260
ttattttccc	agtctcttgg	ccatgatgat	atcttatgat	taaaaacaaa	ttaaattttta	1320
aaacacactga	agatatatta	gaagaaattg	tgcaccctcc	acaaaacata	caaagttaa	1380
aagtttggat	cttttctca	gcaggtatca	gttgtaaata	atgaatttagg	ggccaaaatg	1440
caaaacgaaa	aatgaagcag	ctacatgtag	ttagtaattt	ctagttgaa	ctgtaattga	1500

15

atattgtggc ttcatatgtt ttatTTATA ttgtactttt ttcattattt atggTTTgga	1560
ctttaataag agaaattcca tagTTTTAA tatcccagaa gtgagacaat ttgaacagtG	1620
tattctagaa aacaatacac taactgaaca gaagtgaatG cttatATATA ttatgatAGC	1680
cttaaacCTT tttcctctaa tgccTTAact gtcaaATAat tataaccTTT taaAGCATAG	1740
gactatAGTC agcatGCTAG actgagAGGT aaACACTGAT gcaATTAGAA caggTACTGA	1800
tgctgtcAGT gTTAACACT atgtttAGCT gtgtttATGC tataAAAGTG caatATTAGA	1860
cactAGCTAG tactGCTGCC tcatgtAACT ccaaAGAAAA caggATTCA ttaAGTGCAT	1920
tgaatgtggc tatttCTCTA agttACTCAT attgtcCTTT gcttGAATGC aatGCCGTGC	1980
agatttatGT ggctGCTATT tttatTTCT gtgcATTACT ttaACACCTT aaAGGGAGAA	2040
gcaaACATTt CCTTCTTCAG ctgactGGCA atggCCCTT aactGCAATA ggaAGAAAAA	2100
aaaaAAAGGTT tGTGTGAAAAA ttggGTGATAA ctggcACTTA agatcgAAAAA gaaATTCTG	2160
tataACTTGAT gcTTAAGAT GCCAAAGCT GCCAAAGCT ctgAAAGACT ttaAGATAGG	2220
cagtaATGCT tactACAATA ctactGAGTT ttGTTAGAGT taacATTGA taataAAACT	2280
tgccTGTtTA atctCAA	2297

<210> 13
<211> 655
<212> DNA
<213> Homo sapien

<400> 13	
caggcagCTG ccaggAGCTC ttccCTGCTC gctCACGCT gctCTCAGAA gctCCGATCC	60
agacacACGC gagggCGCTGT CTTTCAGCA ccacaAGCTC gggCTGAGGA gggAGGGACTC	120
ctggCCGTCC tcctCCTCTT caaATTGGCT tgaatCTGCT ctgACCCCCC acgAGTGCAG	180
cacAGTCTGG gaagAAAGGC gtaaggATGG tgaAGCTGAA cagtaACCCC agcGAGAAGG	240
gaaccaAGCC gcTTCAgTT gaggATGGCT tccAGACCGT ccCTCTCATC actCCCTTGG	300
aggTTAATCA ctTACAGCTG CCTGCTCCAG AAAAGGTGAT tgtGAAGACA agAACCGGAAT	360
atcAGCCGGA acagaAGAAC AAAGGGAAgT tccgggtGCC gaaaATCGCT gaATTACGG	420
tcaccATCCT tgtCAGCCTG gCcCTAGCTT tcCTTGCCTG catCGTGTTC ctggTGGTTT	480
acAAAGCCTT cacCTATTG aaggASCTAA attcGTAGCA mattCTGTGG cagTTTAAA	540
aagtTAAGCT gCTATAGTAA gttACTGGC attCTCAATA cttGAATATG gaACATATGC	600
acagggGAAG gaaATAACAT tgcACTTTAT aaACACTGTA ttgtAAgTGG aaaAT	655

<210> 14
<211> 5636
<212> DNA

<213> Homo sapien

<400> 14

aaactgagat ttcaactgat gacaagggttc agaatctgac tgataccgaa gtggttaaga	60
cgcaagagag gaacaaattt tgaggatga agaggtcaga ggtctaagag gccaaagtgtt	120
ggtgttgcattttcacca caagcatgct gaattcagat agaatgtaga caagagggat	180
ggtgaggaaa acctacggca agcagctcta taatttctgg aaagtgacca ggaggctcg	240
ggatgatggc aggaaggaga gtagaaagtg atttagccag aagatatgaa cttcaaata	300
ttttgaagca gaaaagagca gaaatggttt ggaagtagca atcaggacca aagagaccac	360
ccaaactcaa tagcccaacc tcttagcctt agggtacgca gaatctgaaa agtgtatct	420
caattttgga gctctacagc agtaaatctg caagtaacca cctgttagcta atgtccaggg	480
attaaaaaaaaa agataatgaa aatgttttg cccagggaaac attaatttca ggcgtatcta	540
gaatgcagtt ctgcattat acgtactggg ttaccagtca ttgcaaagtc atggccctt	600
gcctctcagc tcagttcccc ttctgaagat aaaaacattt gcctatgtgt ccagggaaagc	660
tgtgaggaca aaaaaccaag caactttac aaggatcat aaaaacctac ctaacaactt	720
gctaattaaa acctgatttt taatttgcattt tattgagctt aataccattt cttaaatgta	780
tgtgaatact gagatttta taaaggaatt agtacctt agggaaataa ttatcactaa	840
aagaaataat atcgcaattt gaataaaagt aagtcgctt aatcatagga aacatTTTAA	900
gtgaaggcgt ttgtttaaa tgtattctaa cctagcaatg tagaaagcag gcaaacattt	960
aaaaaaaaattt taaccagttt ctaaaacata gtttgagct cagattctgg ggaaatgatt	1020
aacacaccct acatccaagg tctccccc tttctagaaa gaaagcatct ttaaacatac	1080
atattcatca gaaatacaaa tatttgcattt cagtatact aatttccagg caaacattt	1140
aattgcagtc aatgtatttag attctaccag gtttaattt ggatcggtaa tacgggTTG	1200
attagggttc taggcctaattat tataggcattt tagtcttctt ttaggattat gcaccatctg	1260
ttatTTAAA ttgaccattt agggggttcc caagggttct cttcgTTTGT gttatcaaac	1320
gttaggttta ggattcttgc ggggtggggg atccaaagcc agagacgggt tcaacaaaaa	1380
tgcaagcgcg aatttgcattt gcgtctgaac gcactgttca aattaaacaa tttagacatg	1440
cccccaactat gacactaaga agtgaatggt atagtaact ttgtcacaatttcaagggt	1500
aatttaagtgc cccgatggta gaggtctggc ttccctggg tttctggaaac aaaagaagcg	1560
ttcgcgagga gaggggttaac tccccccctt cccctccca aagtaaatca aatcaaggaa	1620
tatgagtgc tgcagacaag cctcgcttctt tttttcccccttcaagggttca gctttgggg	1680
aaggcaaggc tgcggctact cttggagctt cagtgccccggg gggggaaaggccagcc	1740

17

aagggtcctc	acactggcgt	ggaattcgcc	gcgttcgtag	gcgatcgacc	ccagagacga	1800
aagctgcttc	tcaagctggg	ggagggagag	gaaacggcgc	acaaaagcag	tacgacctgt	1860
cccttatcgg	cgtctaaggg	gaagggtgga	gaaaacgaaa	acagaagcgg	gccgggagcc	1920
tcggctccc	ccccagcgcc	ttttaaactg	cgtttctacc	tcctctcgct	cagcgccg	1980
gctaattggaa	cccgcgcgag	ccgtctcgcc	aatcacccgc	gcgttcctc	ccgtcgcccg	2040
ccaatggcgg	cgcgcgttct	tggggcgtgg	gcgaagcagg	ctgctcgct	cctgcctgta	2100
gtgtgtggc	tggggtttgt	gctgagttcc	acttggccg	cagtggttc	gtagttcgcc	2160
tctgggtct	tttgtgtccg	ggtctggctt	ggctttgtgt	ccgcgagttt	ttgttccgct	2220
ccgcagcgct	tttcccgggc	aggagccgtg	aggctcggag	gcggcagcgc	ggtccccggc	2280
caggagcaag	cgcgcggcgc	tgagcggcgg	cgccaaaggc	tgtgggagg	gggcttcgca	2340
gatccccgag	atgcccggagt	tcctggaaaga	cccctcggtc	ctgacaaaag	acaagttgaa	2400
gagttagttg	gtcgccaaca	atgtgacgct	gcggccgggg	gagcagcgc	aagacgtgta	2460
cgtccagctc	tacctgcagc	acctcacggc	tcgcaaccgg	ccgcccgtcc	ccgcggcac	2520
caacagcaag	ggggcccccgg	acttctccag	tgacgaagag	cgcgagccca	ccccggtcct	2580
cggctctggg	gccgcgcggc	cgggcccggag	ccgagcagcc	gtcggcagga	aagccacaaa	2640
aaaaactgat	aaacccagac	aagaagataa	agatgatcta	gatgtAACAG	agctcactaa	2700
tgaagatctt	ttggatcagc	tttgtgaaata	cggagtgaat	cctggtccta	ttgtgggAAC	2760
aaccaggaag	ctatatgaga	aaaagctttt	gaaactgagg	gaacaaggaa	cagaatcaag	2820
atcttctact	cctctgccaa	caatttcttc	ttcagcagaa	aatacaaggc	agaatggaaag	2880
taatgattct	gacagataca	gtgacaatga	agaagactct	aaaatagagc	tcaagcttga	2940
gaagagagaa	ccactaaagg	gcagagcaa	gactccagta	acactcaagc	aaagaagagt	3000
tgagcacaat	cagagctatt	ctcaagctgg	aataactgag	actgaatgg	caagtggatc	3060
ttcaaaaaggc	ggacctctgc	aggcattaac	taggaaatct	acaagaggg	caagaagaac	3120
tccaaggaaaa	agggtggaaa	cttcagaaca	ttttcgtata	gatggtccag	taatttcaga	3180
gagtactccc	atagctgaaa	ctataatggc	ttcaagcaac	gaatccttag	ttgtcaatag	3240
ggtgactgga	aatttcaagc	atgcacatctcc	tattctgcca	atcactgaat	tctcagacat	3300
acccagaaga	gcaccaaaga	aaccattgac	aagagctgaa	gtgggagaaaa	aaacagagga	3360
aagaagagta	gaaagggata	ttcttaagga	aatgttcccc	tatgaagcat	ctacaccaac	3420
aggaatttagt	gctagttgcc	gcagaccaat	caaagggct	gcaggccggc	cattagaact	3480
cagtgatttc	aggatggagg	agtcttttc	atctaaatat	gttcctaagt	atgttccctt	3540
ggcagatgtc	aagtcaagaaa	agacaaaaaa	gggacgctcc	attcccgat	ggataaaaat	3600

tttgctgttt gttgttgtgg cagttttttt gtttttggtc tatcaagcta tggaaaccaa	3660
ccaagtaaat cccttctcta attttcttca tggtgaccct agaaaatcca actgaatggt	3720
atctcttgg cacgttcaac ttggctcct attttcaata actgttggaa aacatttgc	3780
tacacttgtt gactccaaga actaaaaata atgtgatttc gcctcaataa atgttagtatt	3840
tcattgaaaa gcaaacaaaa tatataaaa tggacttcat taaaatgttt ttgaactttg	3900
gacttagttagg agatcacttt gtgccatatg aataatctt tttagctctg gaacttttg	3960
taggcttat ttttttaatg tggcatctt atttcatttt tgaaaaaatg tataatgttt	4020
ttgtgtatcc gggaaacgaa gggtaaca tggtagtata atgtgaagct acacattaa	4080
atacttagaa ttcttacaga aaagattttt agaatttattc tctgctgaat aaaaactgca	4140
aatatgtgaa acataatgaa attcagtaag agaaaaagta acttggttgt acctttgtt	4200
actgcaacaa agtttgatgg tggttatgag gaaaagtaca gcaataatct cttctgtAAC	4260
cattttat agtaatgtt tgtagccct atcataactca ctttttaaga cacagtatca	4320
tgaaagtccct atttcagtaa gacccattta catacagtag atttttagca gagatcttt	4380
agtgtaacat acatattttt gagaattgtt ggctagctgt acatgttttggaaatgtt	4440
tagctagcta taaggctata attggaaatt tggatattttt atttacagca aaacatttat	4500
tcagtcattcc agtttgctac caaaatatgt tttagataag tggatgtatg tttgtttaga	4560
agtttagaaat tgtaaacact ggtcttatgt ttcatggaa ttcatatttgcattgttttgc	4620
ttaccagaaa caaatcctcc cgggttcaag caattttccct gcctcgccag agacggggtt	4680
tcaccatgtt gcccaggctg gtctcgacc cctgacctca agtgcacgc ccacccgc	4740
ttccccaaagt gctgggatta caggtgtgat ccactgcacc cggccggcat tatgatTTT	4800
tgtactcttgc aaatggttat ctttggat gatTTTTTT tttaagctga aacttacctc	4860
atgaataact tgattaaagt agtaggtgat taaaatttca atagaatcaa atgagacaaa	4920
aattttaaac tgactcattt gagtttcaac ttacagtca ttgaccataa agcacactaa	4980
aaatgttaagt tattttaaa tacatctgaa ataaaaatac ttactaaaaa ggaagaagcc	5040
gaagatgtat atttagacca gcacacaatt ttgatttcaa ttgccttat tctaataattt	5100
agcttttgcatac acatTTTCAc gtactttgca attgagacca gaaagacttg	5160
taggtctttc tgcagaatga gtgggtcctt gcaaaagttagt tggaaactt actccttagat	5220
cagaaatgtt tgcctctcg agtaaaatgt ttcttcaga tgagccatag agggggcacc	5280
ttttactcaa ctttctttg ttttggaaact ttgtttccca tactgttttc agcctttgt	5340
ttataatttag aaattgttagt aagcttcatt tagtgtttaa aaatgtgggg agataaatca	5400

19

gacttaacat gtagttaaga tcaattcact taaaagtatg gtccaaatag caaaaatagg	5460
accaggtgaa acatgttagtc atttttaaa aacatgtact tggcttttg tgtgtgtctg	5520
ttttattcca ttagaataaa tgtgtccttg atgtaaatgc aaagcattc ttcctgatta	5580
aattgttagat gtagacttta caatataatt caataataaaa aagtaattaa cctcta	5636

<210> 15
<211> 2886
<212> DNA
<213> Homo sapien

<400> 15	
gagccggcta ctggggcg actttcaaa acagcgaaaa caaaacaaat cggggacctt	60
taaaaggcgt aatgagacca gaaacgatct cttccgccc ctctgttcc ccccggtccc	120
caacgcagat caatcgcgga ataagcccga cgcccgatt ccgtctccg ccctagcgcc	180
aggcgggagg actggctcgga caaagccaag gagagctagg gaggccgcga gagaggctcg	240
agacggcagc tttagggcgactctttt taaagtccgt ggaggaagtg caggatccct	300
ccgcggggag tcacgtgccc cggcccttg gggcgctcga aactcttaac aaaaacaagg	360
ggctcgggga ggttccgct gagggcgccgg gggtgccggcg gtggctggc cttccgcggc	420
cggcggtcgcc cggccggcgga gggtggcgccgc gggggagcgg ggtggagct gggcgaccc	480
ttgctggagg tactggcctc agccctttct cccggttccc caccctctt acccccagat	540
tacattctct gtgtgggtgc ttactgcag atgaaggatt tggggcaga gcacttggca	600
ggtcatgaag gggtccaact tctcggttg ttgaacgtct acctggaaaca agaagagaga	660
ttccaaacctc gagaaaaagg gctgagttt attgaggcta ccccgagaa tgataacact	720
ttgtgtccag gattgagaaa tgccaaagtt gaagattaa ggagtttagc caactttttt	780
ggatctgca ctgaaacttt tgcctggct gtcaatattt tggacaggtt cttggctttt	840
atgaaggtga aacctaaaca tttgtctgc attggagtct gttcttttt gctggctgct	900
agaatagttt aagaagactg caatattcca tccactcatg atgtgatccg gattagtcag	960
tgtaaatgtt ctgcttctga cataaaacgg atggaaaaaa taatttcaga aaaattgcac	1020
tatgaattgg aagctactac tgccttaaac ttttgcact tataccatac tattatactt	1080
tgtcatactt cagaaaggaa agaaatactg agccttgata aactagaagc tcagctgaaa	1140
gcttgcaact gccgactcat ctttcaaaa gcaaaaccat ctgtttagc cttgtgcctt	1200
ctcaatttgg aagtggaaac tttgaaatct gttgaattac tggaaattct cttgctagtt	1260
aaaaaacatt ccaagattaa tgacactgag ttcttctact ggagagagtt ggtttctaaa	1320
tgcctagccg agtattcttc tcctgaatgt tgcaaaccag atcttaagaa gttggtttgg	1380

20

atcgtttcaa ggccacacgc ccagaacctc cacaacagct actatagtgt tcctgagctg 1440
 ccaacgatac ctgagggggg ttgtttgtat gaaagtgaaa gtgaggactc ttgtgaagat 1500
 atgagttgtg gagaggagag tctcagcagc tctcctccca gtgatcaaga gtgcaccc 1560
 ttttcaact tcaaagtggc acaaacaactg tgctttccat cttagaaatc tgattgtct 1620
 gtcagaattt atatttacag gtttcaaaggc aataaatggg ggaataggtt gttcctgg 1680
 ttagccccca tctagtcagg aattaatata ctgaaatacc taccttctat ttgttattca 1740
 gatcagatct ggcctatttt catatttata ctaagccatc aaatgggta gtgccttta 1800
 aaccattaac agtactttag acattggcac ttatTTTC tcgtagatct ttagctactt 1860
 tggggaggag ggaaggtgt gataccttca atttgttact tttcaagatt tttaaaaata 1920
 actagtgttag cttatcttaa acattttata aaacctttag atgtcttaa gcagattgga 1980
 agtatgcaag tgcttcctta gcagggacag tggataatcc ttaatggttt atcatagatt 2040
 tcaccctccc cccttctcag aagagtgagt atgctcttaa atgtcaaaca cattttgtt 2100
 gttttttttt ttaaatgatc agtgtctatt tgatgtgatc cagatctt aaatttggga 2160
 attataatat tgacatttct gtgattttta tataatgtat gtcttaattt agatttctgt 2220
 taaggcagaa ataatttaggc tagggctctt agtttcatt cctattgccc aagtattgtc 2280
 aaactatggt attatTTTA tgttacttta aaaatccata atctgctagt tttgcatgt 2340
 cttatatgaa aacagtgcag taagttgaaa actcagttatc tatggaattt ataaatgtt 2400
 atctggtgta gtatattttta tcgcattttc ttatattaaa aaatgtctgc atgattacat 2460
 ttatTTTCT ttgtaattta catttcagaa tagtgttattt cttatgggt gc当地gattt 2520
 aatatgaaga acccgagtgt ttgttagtatt atagttttaa gcaaattctgt gtgggtatc 2580
 agccataaga atggggctta tataaactct gtacatgtaa gattttgtac agagaatttt 2640
 taactttata aattgtatata gaacatgtaa atctttaaa atgtacataa aataactgtat 2700
 ttttttacct tgtgtgtat agtctagtca ttgcattgtaa atataattttt ttatgtattc 2760
 tgttagtataa atcatacatt gatgacttac atttttactg gtaagtcaac atccgttgaa 2820
 tgTTTCTGA agtggctctt tttgaagtgtaa taatagattt gtaattcaaaa taaaattttt 2880
 aatgaa
2886

<210> 16
 <211> 5374
 <212> DNA
 <213> Homo sapien

<400> 16
 tgatacatca ctataggcactgggttcc tagatgtgtc tcgagcggcs scagtgtat60

21

ggatdgccgc	gccccggagg	tactgcttct	gacataaaac	ggatggaaaa	aataattca	120
aaaaaattgc	actatgaatt	ggaagctact	actgccttaa	acttttgca	cttataccat	180
actattatac	tttgtcatac	ttcagaaaagg	aaagaaaatac	tgagcctgta	taaactagaa	240
gctcagctga	aagcttgcaa	ctgccgactc	atctttcaa	aagcaaaacc	atctgtatta	300
gccttgcc	ttctcaattt	ggaagtggaa	actttgaaat	ctgttgaatt	actggaaatt	360
ctcttgctag	ttaaaaaaca	ttccaagatt	aatgacactg	agttcttcta	ctggagagag	420
ttggttctta	aatgccttagc	cgagtattct	tctcctgaat	gttgc当地	agatcttaag	480
aagttggttt	ggatcgtttc	aaggcgcaca	gcccagaacc	tccacaacag	ctactatagt	540
tttccctgagc	tgccaaacgat	acctgagggg	ggttgggggg	atgaaagtga	aagtgaggac	600
tcttgtaag	atatgagttt	tggagaggag	agtctcagca	gctctccccc	cagtgtatcaa	660
gagtgcaccc	tcttttcaa	cttcaaagtg	gcacaaacac	tgtgcttcc	atcttagaaa	720
tctgattgtt	ctgtcagaat	ttatatttac	aggtttcaaa	gcaataaaatg	ggggaaatagg	780
tagttccctg	gttagcccc	catctagtca	ggaattaata	tactggaata	cctacccct	840
atttgttatt	cagatcagat	ctggcctatt	ttcatattta	tcctaagcca	tcaaattgggg	900
tagtgccct	taaaccatta	acagtacttt	agacattggc	actttatttt	tctcgtagat	960
cttttagctac	tttggggagg	agggaaagggtg	ctgatacctt	caatttgtt	ctttcaaga	1020
tttttaaaaa	taactagtgt	agcttatctt	aaacattttt	taaaacccctc	agatgtcttt	1080
aagcagattt	gaagtatgca	agtgc当地	tagcaggac	agtggataat	ccttaatgg	1140
ttatcataga	ttcacccctc	ccccctctc	agaagagtga	gtatgcttt	aaatgtcaaa	1200
cacatttttt	ttgtttgtt	ttttaatga	tcagtgtcta	tttgatgtga	tgcagatctt	1260
ataaaatttg	gaattataat	attgacattt	ctgtgatttt	tatatatgt	atgtcttaat	1320
tgagatttct	gttaaggcag	aaataattag	gttagggctc	ttagtttca	ttccatttgc	1380
ccaagtattt	tcaaaactatg	gtattatttt	aatgttactt	taaaaatcca	taatctgcta	1440
gttttgcatt	tacttatatg	aaaacagtgc	agtaagttga	aaactcagta	tctatgg	1500
tgataaatgt	tgatctggtg	tagtatattt	tatgcattt	tcttatatta	aaaaatgtct	1560
gcatgattac	attttatttc	cttgttaatt	tacatttcag	aatagtgtat	tgctatatgg	1620
gtgccaagat	tgaatatgaa	gaacccgagt	gtttgttagt	ttatagtttt	aagcaaattct	1680
gtgtggtgat	acagccataa	gaatggggct	tatataaact	ctgtacatgt	aagattttgt	1740
acagagaatt	tttaacttta	taaattgtat	atgaacatgt	aaatcttttta	aaatgtacat	1800
aaaatactgt	atttttttac	cttgtgtgt	atagtctagt	cattgcattt	aaatataatt	1860
tattatgtat	tctgttagtat	aaatcataca	ttgtgactt	acatttttac	tggtaaagtca	1920

acatccgttg gatgtttct gaagtggctc ttttgaagt gataatagat tgtaattcaa	1980
aataaaatta ttaatgaatt ctccttgtt gggatcacat cttaatttt aatctgttaa	2040
aagttcttga tgtatTTAA tgagaagact ttaggtgagg ctacagtgtat tccagagtga	2100
gccttctaacc tggcttagcag aagttctcta ggttggcat ctgtgcctt gagatactga	2160
aagagaatct gtcatttgac aattgacctc ttgtggat ggactcatta agtatgctct	2220
cagagactgg tatattacca gaatgcctat taatTTCA tgagaggcaa caggtattaa	2280
gtagaacaga atgctcagg tggcagatta gaacgatctt tcaggagaca aagcaagttt	2340
taatcagttt tttggtaat aagtatgggg tggcgctgt gatagggccc cgccagcttc	2400
tggctttgtt ggacctcaaa agtacggat ggtttgcaa gtgggtgtcc tttccccctgc	2460
cccaccccaa tagttcccc atctgtctag ttgatTTT gtagacctt gtttctcta	2520
gttagaaaat caggtacact gaatatggtt ttcatgtaac acctttctc tggagatagg	2580
ggtagttttt cctacccttc tagtggagaa tcctacttga ggatgacctt tcctcttta	2640
ctaaataata ttagtaaata gtggcaata tattctgtt tcagatTTG atttggtag	2700
atgtaaaagt tgTTTggggc ttaccaaatac tcaagactct ctttagctcc tgcaggattt	2760
tattgtttt cttaactggat atttttctg ggtaagcattt tttgtggctt catctttcc	2820
ccctgtggtt ttcaagtgtat tttagtcgaga cctctctgt gacgttgcac cctgtttatt	2880
cacatggcct gccatgccac ttggaggTTT ctgattactc ccaaaccgtc tggttttta	2940
tgtctttctc agcgaataat tccatctgtt catgtggaa acttaggtga tatgtctatc	3000
tccttttgcg ttttatggaa ggtcaccaggc ctctatcatt tttatgtt cgtttacact	3060
gtttatatct ctctgtcccc ctttttctg ccattggcat gtttagacc tttttttttt	3120
atcagcagag gtactgtaat atatttgc tccctcagct tccaggctt ctcctggct	3180
ctgccttcct atctacatat cttttaaaaaaa taaaatttttta actatctcct gaaaaattgt	3240
ttagtaggtc acgcacaatac aggagaaaaa tctattcatg acatacaagt ctctgtctaa	3300
tctgaacact gcacctgtct ctggcctttt tttcttgc tttccttagac cttaaaaaat	3360
gtgtattgag aaagaactct gtttagctata cagaagatga actgggcaat atagagtagc	3420
agcatggaga ccagtctgac tgaactaagg cagtgaaagt gtggatgagg aagagaggtg	3480
aaaattgaga agcgctatcc tttctttttt ggcattatttta ggaggctcac agacaagtcc	3540
aggagcctgg ttataccctc ctgtgccatt caaccagggtg gttttccat gactgtgtat	3600
aataaaattt agaagccccctt gcccTTTca gagcagaggg tgaggagaaa gctaccattt	3660
tgtcctcatc cttacccccc ttgacttggc gagagatttgc acctttcagg ttttgcattt	3720

23

gtcattttctt aggatgtggt gcacgcactt tgctgttgcg catggtaag tattgtgcct	3780
aggtcctggg tcttcatctg tttggctctg ctactgttcc ctcctcccag gaagtgtggt	3840
tagacaata atgtgttttta attacctgtc acactcagga ttaatacata ctcaggttaa	3900
ctgttagagag gcattggcctt cagaacactc ctcgtgacaa ttttaaccat tttctttgtc	3960
tagagtctgc ctttttctttt ttacaattt cttttatttc aacacttaggt ttcaatatgg	4020
tgttcctgct acctcccacc tccctcctcc ctcatcacac atgcaaatttgc ttagcttatt	4080
gagacaaccc acttagattc atatatggac aaggacaagg tattttgcattt ttgttactgg	4140
aattcagttt tcctaactat ttactaccag aaatggtcaa taacttactt tgtgttttagc	4200
aaatcaaaattt gtgtgataga tagttttccca gtagtgcgc cagtcagttt ttccatccct	4260
gtgcctacat gctgctttcc ccgtccacaa gtggagtctg tttctcttga gttttggctg	4320
gccttatgaa tggctttgtt tactgaagtg cagcagaaga aatttagtat atgtccaagg	4380
ctaggcttta agagactggc agctttcctt ttatcctttt tggaaagctt ccaccatgtt	4440
gcaaagaagc ttagctggat tactgaaaaga tggaggccat tggagggaga gactcttgag	4500
gatgagagat tatcttggat gttccagcct taagctccca gctgaatgtg ggtgtatcct	4560
cagctacacc acagaaaaca gaggaactac tcaatcgatc ccaatcaacc cacagactca	4620
ctagaaataaa caaatttattt ttttaagcca cgaggtttg ggggagggtt gttaaacagt	4680
aatagataag tgagacagat tgcttggat ttatggtcaa atgggtatta tctctggat	4740
gattacaggt gatgtttttt ttaagttatg cctatctgtt gtttcctttt tttcctaaaa	4800
ttgatttggaa ttatttgtt attaacagaa taaaagatga actttaaaac acacacgctt	4860
gttatatgtt tcctctaattt aaaattcatg gctctcacca caccttagca tcaagttccca	4920
acttcgtact gcccgtttaga agaccagct tgatttggat ccggctccct tttcagcctt	4980
gtttcatggc atccacatcc acgtattcc caggccact acatctgaga tgagtcagag	5040
acccctctta ggggcctgtt ccctactccc aaacatggaa attaaaaaaaaaa aaatcgtag	5100
ttcttccaag agaaattcca ggcattctggc tagccctgag aagtaagaga gaaatgtgat	5160
aagcaacaaa tagcggctca aaacaatagc caagtaagtt agaatcatgg gatgtttggat	5220
tccctatag aaactacaga taacatctta atatatacc ctgagttgtt ttccagaaac	5280
ctgaacccctt agcaaatttggat tgcgttagca catagaccc agataagagg gagctgagga	5340
ctgaactctg accaccgttc tttgttctaa attt	5374

<210> 17
 <211> 663
 <212> DNA
 <213> Homo sapien

<400> 17
ccagaaaaaca acagaatgaa catcatcatg aatacatgaa tcggctgtga tgtgtgaact 60
gctaagggcc aaatgaacgt ttgcagagca gtgggcacaa tgtttacaat gtatgtgtat 120
gtcactttcg gtacctgtga atgcattttt acgtgctgaa cccgaaaaaaaaa agtgcctttc 180
cataaggact gcaatagaga gggcaattta ccctgggtggt acacggAACCC tagattcaact 240
cctgccatgc cttgccaata gtaagctgca gggtggaaca agaaatcaact tgctctgggg 300
ggaagggagg gggaaatggg tgtgtcagct ggtagatac aaaccctgaa aagagaatcc 360
atgtgctgct ggcaggcaac attttttaaa gctcttcag aaaccctcat atttggggtt 420
tcttttcagg aaacatttcct gtggaggggaa aacgaatatg aagataattt tcagctaatt 480
atctgggtga cccagaatcg tgttatatggc tataggatag acttcttaat aatggcaagt 540
gacgtggccc tggggaaagg tgctttatgt accgtgtgtg cgtgtatgtg tgtgtatcta 600
tacaagtttgc tcagctttgg catgactgtt tgtttgtctc gaaaaccaat aaactcaaag 660
ttt 663

<210> 18
<211> 2162
<212> DNA
<213> Homo sapien

25

tggctgtcaa	gtgtgccacc	atcacccctg	atgaggccccg	tgtggaaagag	ttcaagctga	900
agaagatgtg	aaaaagtccc	aatggaacta	tccggAACAT	cctgggggggg	actgtttccc	960
gggagcccat	catctgcAAA	aacatcccac	gcctAGTCCC	tggctggACC	aagcccatca	1020
ccattggcag	gcacgcccatt	ggcgaccagt	acaaggccac	agactttgtg	gcagaccggg	1080
ccggcacttt	caaaaatggtc	ttcaccccaa	aagatggcag	tggtgtcaag	gagtgggaag	1140
tgtacaactt	ccccgcaggc	ggcgtggca	tgggcATGTA	caacaccgac	gagtccatct	1200
caggtttgc	gcacagctgc	ttccAGTATG	ccatccagaa	gaaatggccg	ctgtacatga	1260
gcaccaagaa	caccatactg	aaAGCCTACG	atgggcgttt	caaggacatc	ttccaggaga	1320
tctttgacaa	gcactataag	accgacttcg	acaAGAATAA	gatctggat	gagcacccgc	1380
tcatttatGA	catggtggtc	caggtcctca	agtcttcggg	tggctttgtg	tgggcctgca	1440
agaactatga	cggagatgtg	cagtcagaca	tcctggccca	gggctttggc	tcccttggcc	1500
tgtatgacgtc	cgtcctggtc	tgcctgatg	ggaAGACGAT	tgaggctgag	gccgctcatg	1560
ggaccgtcac	ccgcccactat	cgggagcacc	agaAGGGCCG	gcccaccAGC	accaACCCCC	1620
tgcCcAGCAT	cttgcctgg	acacgtggcc	tggAGCACCG	ggggAAAGCTG	gatgggaacc	1680
aagacctcat	caggTTTGCC	cagatgctgg	agaAGGTGTG	cgtggagacg	gtggagAGTG	1740
gagccatgac	caaggACCTG	gcgggctgca	ttcacggcct	cagcaatgtg	aagctgaacg	1800
agcacttcct	gaacACCACG	gacttcctcg	acaccatcaa	gagcaACCTG	gacAGAGCCC	1860
tgggcaggca	gtagggggag	gcGCCACCCA	tggctgcAGT	ggagggggcca	gggctgagcc	1920
ggcgggtcct	cctgagcgcg	gcagAGGGTG	agcctcacag	cccccTCTCTG	gaggcTTTC	1980
tagggatgt	tttttataa	gccagatgtt	tttaaaAGCA	tatgtgtgtt	tcccttcATG	2040
gtgacgtgag	gcaggAGCAG	tgcgtttac	ctcAGCCAGT	cagtatgttt	tgcataACTGT	2100
aatttatatt	gcccttggaa	cacatggtgc	cataTTAGC	tactaaaaAG	ctcttcacaa	2160
aa						2162

<210> 19
<211> 1527
<212> DNA
<213> Homo sapien

<400> 19	ggcACGAGAA	ttggcAGACT	ccAGAGCCCA	cacATTGCA	CTCTAGACTC	tactgcTTTC	60
	ctcatGAAGA	atTTTAGGAC	ccccgtctgg	ctgtgtgtt	gcttgggggtt	caaattctgg	120
	ttgaaAGATG	gcggcTGCAG	tgggACCACT	attatCTCTG	tcctCACAGA	gttcaAGCTG	180
	aagaAGATGT	ggAAAAGTCC	caatGGAACT	atccGGAACA	tcctgggggg	gactgtttc	240

26

cgggagccca	tcatctgcaa	aaacatccca	cgcctagtcc	ctggctggac	caagccccatc	300
accattggca	ggcacgcccc	tggcgaccag	gtaggccagg	gtggagaggg	gatccactga	360
cctgggcacc	ccccgactgg	agctcctcgc	ctagccatcc	tcttgtctct	gcagtacaag	420
gccacagact	ttgtggcaga	ccggggccggc	actttcaaaa	tggtcttcac	cccaaaaagat	480
ggcagtggtg	tcaaggagtg	ggaagtgtac	aacttccccg	caggcggcgt	gggcataggc	540
atgtacaaca	ccgacgagtc	catctcaggt	tttgcgcaca	gctgcttcca	gtatgccatc	600
cagaagaaat	ggccgctgt	catgagcacc	aagaacacca	tactgaaagc	ctacgatggg	660
cgtttcaagg	acatcttcca	ggagatcttt	gacaaggact	ataagaccga	cttcgacaag	720
aataagatct	ggtatgagca	ccggctcatt	gatgacatgg	tggctcaggt	cctcaagtct	780
tcgggtggct	tttgtgtggc	ctgcaagaac	tatgacggag	atgtgcagtc	agacatcctg	840
gcccagggct	ttggctccct	tggcctgatg	acgtccgtcc	tggctgccc	tgtatgggaag	900
acgattgagg	ctgaggccgc	tcatgggacc	gtcacccgcc	actatcgga	gcaccagaag	960
ggccggccca	ccagcaccaa	ccccatcgcc	agcatcttg	cctggacacg	tggcctggag	1020
caccggggga	agctggatgg	gaaccaagac	ctcatcaggt	ttgcccagat	gctggagaag	1080
gtgtgcgtgg	agacggtgga	gagtggagcc	atgaccaagg	acctggcggg	ctgcattcac	1140
ggcctcagca	atgtgaagct	gaacgagcac	ttcctgaaca	ccacggactt	cctcgacacc	1200
atcaagagca	acctggacag	agccctgggc	aggcagtagg	gggaggcgcc	acccatggct	1260
gcagtggagg	ggccagggct	gagccggcgg	gtcctcctga	gcgcggcaga	gggtgagcct	1320
cacagccct	ctctggaggc	ctttctaggg	gatgttttt	tataagccag	atgttttaa	1380
aagcatatgt	gtgtttcccc	tcatggtgcac	gtgaggcagg	agcagtgcgt	tttacctcag	1440
ccagtcagta	tgtttgcat	actgtaaattt	atattgcct	tggaacacat	ggtgcctat	1500
ttagctacta	aaaagcttt	cacaaaaa				1527

<210> 20

<211> 1364

<212> DNA

<213> Homo sapien

<400> 20

ccaaaaaaaaaa	aaaaaaggcg	gtgtttaca	aagcaaagtt	gagagggaga	ggctgggcca	60
gcagaaacat	cgtgtgcact	gcacggaggc	tggtgtaaa	cagtcgcgtg	ggcggcgggg	120
taccgttcct	ggagagctgg	gccttgcct	gggaggtggg	aggttgccgg	caatcgccag	180
gctagggcac	cacgccaggg	ccctgtctct	ccccctgcag	tccatctcag	gttttgcgca	240
cagctgcttc	cagtatgcca	tccagaagaa	atggccgctg	tacatgagca	ccaagaacac	300

27

catactgaaa gcctacgatg ggcgttcaa ggacatctc caggagatct ttgacaagta	360
aagcctcatc catgtactct gtggccttc ttcccttccc cccatgctgt tcccatccta	420
ccctgggaag gtcgttatta gagtgattt ggctagctc cgaggctcag ggagggatcc	480
ccaacctgtc agccttctgc cctctcccc taacagacct ttttactccc aggcactata	540
agaccgactt cgacaagaat aagatctggt atgagcaccc gctcattgat gacatggtgg	600
ctcaggtcct caagtcttcg ggtggcttg tgtggcctg caagaactat gacggagatg	660
tgcagtcaga catcctggcc cagggottt gctcccttgg cctgatgacg tccgtcctgg	720
tctgccctga tggaaagacg attgaggctg aggccgctca tgggaccgtc accccgccact	780
atcgggagca ccagaagggc cgccccacca gcaccaaccc catcgccagc atcttgcct	840
ggacacgtgg cctggagcac cggggaaagc tggatggaa ccaagacctc atcaggtttgc	900
cccagatgct ggagaagggtg tgcgtggaga cgggtggagag tggagccatg accaaggacc	960
tggcggtctg cattcacggc ctcagcaatg tgaagctgaa cgagcaccc ctgaacacca	1020
cggacttcct cgacaccatc aagagcaacc tggacagagc cctggcagg ctaggggg	1080
aggcgccacc catggctgca gtggaggggc cagggcttag ccggcgggtc ctcctgagcg	1140
cggcagaggg tgagcctcac agccctctc tggaggcctt tctagggat gtttttttat	1200
aagccagatg ttttaaaag catatgtgtg tttccctca tggtagcgtg aggcaggagc	1260
agtgcgtttt acctcagcca gtcagttatgt tttgcataact gtaattata ttgcccttgg	1320
aacacatggt gccatattta gctactaaaa agctttcac aaaa	1364

<210> 21
<211> 897
<212> DNA
<213> Homo sapien

<400> 21	
accctccagt gcccctccagc cctgtgtgg gcccctggaga cccacaggag ggtgaagaga	60
cctggAACAG tccctgtcct cccagtgca gctggggag gctgagttaga gcccacgaact	120
atggcagcta caatattggg ttgttagaggg cagcaggct cagctgggtg gccccaggag	180
aggcgaggcc ctgagagaaa ggctttctac cctccaggct ttggctcct tggcctgatg	240
acgtccgtcc tggctgtccc tggatggaaag acgattgagg ctgaggccgc tcattggacc	300
gtcacccggcc actatcgaaa gcaccagaag ggccggccca ccagcaccaa ccccatcgcc	360
agcatctttg cctggacacg tggcctggag caccggggaa agctggatgg gaaccaagac	420
ctcatcaggat ttgcccagat gctggagaag gtgtgcgtgg agacgggtgaa gagttggagcc	480
atgaccaagg acctggcggg ctgcattcac ggcctcagca atgtgaagct gaacgagcac	540

28

ttcctgaaca ccacggactt cctcgacacc atcaagagca acctggacag agccctgggc	600
aggcagtagg gggaggcgcc acccatggct gcagtggagg ggccagggtc gagccggcgg	660
gtcctcctga ggcggcaga gggtgagcct cacagcccct ctctggaggc ctttctaggg	720
gatgttttt tataagccag atgttttaa aagcatatgt gtgttcccc tcattgtgac	780
gtgaggcagg agcagtgcgt tttaccttag ccagtcagta tgtttgcatt actgttaattt	840
atattgcctt tggAACACAT ggtGCCATAT TTAGCTACTA AAAAGCTCTT CACAAAA	897

<210> 22

<211> 1548

<212> DNA

<213> Homo sapien

<400> 22

tgcccgcgcg gccagcgccc gccaggccca gcgttagcgt tagcccgccg ccaggcagcc	60
gggaggagcg ggcgcgcgtc ggacctctcc cgccctgctc gttcgctctc cagcttggga	120
tggccggcta cctgcgggtc gtgcgctcgc tctgcagagc ctcaggctcg cggccggcct	180
gggcgcggc ggcctgtaca gcccccacct cgcaagagca gccgcggcgc cactatgccg	240
acaaaaggat caaggtggcg aagccctgtgg tggagatgga tggatgatgag atgaccgcgt	300
ttatctggca gttcatcaag gagaagtgtg aagctgaacg agcacttcct gaacaccacg	360
gacttcctcg acaccatcaa gagcaacctg gacagagccc tgggcaggca gttaggggag	420
gcgcaccca tggctgcagt ggaggggcca gggctgagcc ggcgggtcct cctgagcg	480
gcagagggtg agcctcacag cccctctctg gaggccttc tagggatgt tttttataa	540
gccagatgtt ttAAAAGCA tatgtgtgtt tccctcatg gtgacgtgag gcaggagcag	600
tgcgtttac ctcagccagt cagtatgttt tgcatatgt aatttatatt gcccggaa	660
cacatggtgc catatttgc tactaaaaag ctttcacaa aattgtctgc tgggtttgtc	720
cctgagggga ggaggtatgt ggaccctgag gcagaggccc tgctagagct ggcagggtcc	780
cctggggcag accagagcac ctcaggaagg ggctgccacg gcaggaaagg gaccaggcag	840
ccctgggagc ccgcattcca cagggccca ctgcggagtt ctcggacact cagggcacag	900
gcctgtgggt tccctggaat tttctagcat gatccagttt ctgtgtccag ttctccatc	960
tgagagtcaa tcagttcctg ataggttgtc attgatttt ttcttcgttg gttttacat	1020
tctaaacatc tccaggccac tttcttagcc ttttcttagg tactaaaaag aggtcttacc	1080
cacacctgcc tcacacttct ctttccaag gctgcctgag tttggagggg cttgggtgtg	1140
tgtgaacaag ggcctgcat tgtcttagcc tgcagttccc aggctgggt tcactttcac	1200
catgcattgg caaaactaga aaagtaagct tgtgacaaat tttctcgcc cgggcacagt	1260

29

ggcgacgccc tataatccct gtactttggg aggctgaggt gggtgatca cttgaggcca	1320
ggagttcgag accagcctgg ccaacatggt gaaacccat ctctactcaa aataaaaaaa	1380
ttagccaggc gtggtgatgc gcacctgcag tccagctac tcgggaggct gaggcaggaa	1440
aatggcttga acctgggagg cagaggttgc agtgagccga gactgcacca ctgcactcca	1500
gcctgggtga cagagcaaga ctctgtctca aaaaaaaaaa aaaaaaaaaa	1548

<210> 23
<211> 3393
<212> DNA
<213> Homo sapien

<400> 23	
acactgggtt cgagttccca acctcaggc atctgcccgt ctcagcctcc caaagtgc	60
ggattacagg cgtgagccac cgtgcctggc cgtaaggtat tattttaaa ggtagctca	120
cctaagactt ccgcagctga gggcagtaac aagataggca tgatgcacag agccatgtgg	180
ggattccaga cccctcctgt tgcatagttt cccagttgaa tttgactttt ctccatttat	240
ctcattttt tctggatagg tctacctgca agtcggattt cccaggttat tgttggagat	300
gagcggcagc aataacctt ggtgattggg caggtttag ttagtgcagg tttagtcagc	360
gtttggctca ggcgaatgaa attgtcctat cctggaaactg ctggatgctt tgcaagcagt	420
atatgtttga agtggcaatc atgaggagg atgaagctgt gaagattgtt gaaggccagc	480
ctatacgttata tttatctgaa ttccgtacga tgactctggt tttggctcagc cttagatcc	540
acaggacagc gtggatgtt catttgttc atcttatcca ggaggctgcc ttatacatct	600
ccacagtcat tgagaaagggg ggcggccagc tgagtcggat ctttatgtt gagaaaggct	660
gcattttctt ctgtgttttc ggccttcctg gtgataagaa gccagacgag tgtgcacatg	720
ccctggagag ctcccttcagc atcttcagct tctgctgggaa gaatcttgct aagaccaact	780
gaggaggaag gtggggcaga gaggagcttc tcaggccccca ggggctttc aggccaggatc	840
cctagatttgc ttccatcag tatcaactt ggaccagtat tctgtggcgt ggttggagca	900
gtagcaagac acgaatatac agttattggc ccaaaagtga gtcttgcggc cagaatgata	960
actgcttatac caggtttggt gtcctgtat gaggtacat atctaagatc catgtacat	1020
gtttacaact tcaagaaact cccagagaaa atgatgaaaa acatctccaa cccagggaaag	1080
atataatgtt atcttggcca cagaagatgt ataatgtttt gaaaaagaca tttggcaaga	1140
aagagaaaca aaaatcaccc ttgttagga gtgttaggtg ctccctgtct ctctacagac	1200
tgggagaaag aatttggaaagc cttccaaatg gcacagcaag ggtgtttgca ccagaagaag	1260
ggacaagcag ttctgtatga aggtggaaaa ggctatggaa aaagccagct gttggctgaa	1320

ataaaacttgc	tggcacagaa	agaaggccat	agctaccctt	cacaggtgct	ttggaaaccc	1380
actttattgt	gaggtcctat	gccaggacct	tctctctaag	gacgtgtgc	tcttcatgt	1440
cctacaaaag	gaggaagagg	aaaacagcaa	gtggaaacc	ctctcagcca	atgccatgaa	1500
atccataatg	tatagtattt	ctcctgccaa	ctctgaggaa	ggccaggaac	tttatgtctg	1560
cacagtcaag	gatgtatgt	acttggatac	agtacttctc	ctacccttt	tgaaagaaat	1620
agcagtaagc	caactggatc	aactgagccc	agaggaacag	ttgctggtca	agtgtgctgc	1680
aatcatttgtt	cactccttcc	atatacgattt	gctgcagcac	ctcctgcctg	gctggatata	1740
aaataagcta	cttcaggctct	tgagagctct	tgtggatata	catgtgctct	gctggctga	1800
caagagccaa	gagttcctg	ctgagcccat	attaatgcct	tcctctatcg	acatcattga	1860
tggaacccaaa	gagaagaaga	caaagttaga	tggtgggtca	gcctctcttc	tcaggctaca	1920
agaagaatta	tccttaccac	aaactgaggt	gttggatattt	ggagtgcctc	tgctacgggc	1980
agctgcttgg	gagctctggc	ccaaggaaca	acagatagct	ctgcaccttg	aatgtgcctg	2040
ctttctccaa	gttttggct	gccgctgtgg	gagctgccat	ggaggagact	ttgtccccctt	2100
tcatcatttt	gcagtttgtt	ctactaagaa	ttccaagggg	acctctcgat	tctgtactta	2160
cagagatact	ggctcagtgc	taacacaagt	gatcacagaa	aaattgcagc	tgccttctcc	2220
ccaagaacag	aggaagagtt	cctagatcaa	gtgaagagga	agctggctca	gaccagccct	2280
gagaaagacc	tgttgaccac	aaagccttgt	cactgttaagg	atatcctgaa	gttagtgctc	2340
ttacccctca	cccagcattt	cttggcgttt	ggagaaacca	cctgtgcatt	ttattacctg	2400
ctggaggctg	cggtgcctg	cttggacctg	tcagataatt	atatggctg	tttcaacatg	2460
ggacgtatca	ctttagccaa	aaaattggct	aggaaagccc	ttcgactgct	gaaaaggaat	2520
ttcccttgg	cctggtttgg	tgtccttttc	cagacattcc	tggaaaagta	ttggcattcc	2580
tgtaccctga	gcacaccc	aaacgaccct	agtgagaagt	tgtctctacc	tatgtggagc	2640
tctctcagtt	ctcccaagagt	gtgggcatca	aggacaagt	gctgcactgt	gagcagatgg	2700
ccattcagaa	aagcagttt	tgttggttct	ccagggaggg	gttggggcc	acagctcagc	2760
tcatgcaggg	cctggcctac	accaagctct	gccttggtca	tcttgacttc	tccatcaagc	2820
tgggattgct	gtgtcgcccc	tttagtgagt	gtctgcgttt	cgttcaagtc	tacgagcaca	2880
gccgtttct	aacctctcag	agcaatgtca	tgctgggggt	ccactcctcc	ctggccatgt	2940
gtactgtaag	gagttcttct	ctcaatgtgt	gacctgcct	gtctatcacc	agtgggtatc	3000
tgagcttaag	gcctctgtaa	tgagatgtga	aaagagagaa	ttgatgtccc	tgactaacag	3060
catcagacct	tttgacacact	gcttgaccag	gattggata	aaaggagaat	ttctgcagga	3120
aaataactct	tagaaaagaa	acttaggaat	acagagattt	gacagagttt	ctgatgtcaa	3180

ggagaacaag gatcagaag aaactcaaga tgtatgtatt aaaacaaaag aacaataacc	3240
tgaagggacc atgattctgt tattgtatat aacacaagga aatgccccag atttccttt	3300
aaaagatata atgtacatat taagtatact agccttata gttactgcta tctacatgtt	3360
tatcaaaaata aaagactatt ttttctaaa aca	3393
<210> 24	
<211> 4034	
<212> DNA	
<213> Homo sapien	
<400> 24	
ggtaactcaa tgtgtctta tagattgatg aaggccagcc tatacgat gtatctgaat	60
tccgtacgt gactctgggtt ttggcagcc tagagttcca caggacagcg tggatgttgc	120
atttgtgtca tcttatccag gaggctgcct tatacatctc cacagtcat gagaagggg	180
gcggccagct gagtcggatc tttatgttg agaaaggctg catgttcctc tgtgtttcg	240
gccttcctgg tgataagaag ccagacgagt gtgcacatgc cctggagagc tcctcagca	300
tcttcagctt ctgctggag aatcttgcta agaccaactg aggaggaaga tccttagatt	360
tgttccatc agtatacata atggaccagt attctgtggc gtgggtggag cagtagcaag	420
acacgaatat acagttattg gccaaaaagt gagtcttgcg gccagaatga taactgctta	480
tccaggttg gtgtcctgtg atgaggtAAC atatctaaga tccatgtac ctgottacaa	540
cttcaagaaa ctcccagaga aaatgtgaa aaacatctcc aacccaggg agatatatga	600
atatcttggc cacagaagat gtataatgtt tggaaaaaga cattggcaa gaaagagaaaa	660
caaaaatcac cctttgttag gagtgtagg tgctccctgt ctctctacag actggagaaa	720
agaattggaa gccttccaaa tggcacagca aggggttttgc caccagaaga agggacaagc	780
agttctgtat gaaggtggaa aaggctatgg aaaaagccag ctgtggctg aaataaactt	840
tctggcacag aaagaagggc atagctaccc ttacaggtg ctttgaaac ccactttatt	900
gtgaggtcct atgccaggac cttctctcta aggacgtgtt gcttttcat gtcctacaaa	960
aggaggaaga ggaaaacagc aagtggaaa ccctctcagc caatgccatg aaatccataa	1020
tgtatagtat ttctcctgcc aactctgagg aaggccagga actttatgtc tgcacagtca	1080
aggatgatgt gaacttggat acagtacttc tcctaccctt tttgaaagaa atagcagtaa	1140
gccaactgga tcaactgagc ccagaggaac agttgctggc caagtgtgct gcaatcattg	1200
gtcactcctt ccatacatat ttgctgcagc acctcctgcc tggctggat aaaaataagc	1260
tacttcaggt ctgagagct ctgtggata tacatgtgct ctgctggct gacaagagcc	1320
aagagcttcc tgctgagccc atattaatgc cttcctctat cgacatcatt gatgaaacca	1380

aagagaagaa gacaaagtta gatgggggt cagcctctt tctcaggcta caagaagaat	1440
tatccctacc acaaactgag gtgttgaat ttggagtgcc tctgctacgg gcagctgctt	1500
gggagctctg gcccaaggaa caacagatag ctctgcacct tgaatgtgcc tgcttcctcc	1560
aagtttggc ctgcccgtgt gggagctgcc atggaggaga ctttgtcccc tttcatcatt	1620
ttgcagtttgc ttctactaag aattccaagg ggacctctcg attctgtact tacagagata	1680
ctggctcagt gctaacacaa gtgatcacag aaaaattgca gctgccttct ccccaagaac	1740
agaggaagag ttcttagatc aagtgaagag gaagctggct cagaccagcc ctgagaaaga	1800
cctgttgacc acaaagcctt gtcactgtaa ggatatcctg aagttagtgc tcttaccct	1860
cacccagcat tgcttggctg ttggagaaac cacctgtgca ttttattacc tgctggaggc	1920
tgcggtgtcc tgcttggacc tgtcagataa ttatatggc tgtttcaaca tgggacgtat	1980
cacttagcc aaaaaattgg ctaggaaagc cttcgactg ctgaaaagga atttcccttgc	2040
gacctggttt ggtgtccctt tccagacatt cctggaaaag tattggcatt cctgtaccct	2100
gagccaacctt ccaaacgacc ctagtgagaa gttgtctcta cctatgtgga gctctctcag	2160
ttctcccaaga gtgtggcat caaggacaag tggctgcact gtgagcagat ggccattcag	2220
aaaagcagtt tatgttggtt ctccagggag gggttgttgg ccacagctca gctcatgcag	2280
gccctggcctt acaccaagct ctgccttggc catcttgact tctccatcaa gctgggatttgc	2340
ctgtgtcgcc ctttagtga gtgtctgcgt ttgttcaag tctacgagca cagccgtgtt	2400
ctaacctctc agagcaatgt catgctgggg gtccactctt ccctggccat gtggtaatgt	2460
cttactcaag ggctgtggaa aaggatagac atttatgtca tttaagctgt ctctccccac	2520
cagacaggac tttgaacctt ctctaaccat ctttaaaga ccattcacct cccataccct	2580
cccatcttat tagaagggtt cttgtccctt aacaggtttt ggcctatagg tcaagggtta	2640
cgtttagggt tacatttaac tgctagagta acccatagca aggctgaata taattggct	2700
ccttttaagt ttccctgtat gtgagttgtt agccttggc actttcttagc atcacaattc	2760
tgattgtcca tgaggtctt gagccttaaa gaagtgtatga tttaagcaa aagtcatggt	2820
gggtaagcag cggatattgc tgcaagctgt tactcttttcccttgcgtt gcccaggaat	2880
cacagtggga cctgtttaag cactattctt ccaacgcttgcgtt gggatgttgcgtt aaagaaccaa	2940
tgcctcgcta tttggtgac atggcttgcgtt ccgattccta gaatgccatg tgtaatgtt	3000
acagaaaatg ccagagggtt tcttcatgca tattcctcta gagttcaca gccaaaccct	3060
tgaggtaccc tttctcagc tgcttgcgtt ctaacaccttgc attcacttag ttctacccttgcgtt	3120
tggtgcttctt tctaccaccc tttttttcccttgcgtt gcatctcttc	3180

33

tttactctt	gaatccttg	tttctccacc	tagaaagttt	ctacacct	tatgtatcct	3240
tcccgatatt	attgcata	gttctggact	gggaaaat	actttccacc	tttgcagct	3300
gctacccagt	atcattaaaa	tattaacatt	taagecttgct	caatggac	ctgtatg	3360
gttcagtc	taatttgata	cagctccctc	cagcccttct	gagtctaaaa	cacattccaa	3420
ttccctgtt	ttccaggctt	attttgc	cagtaactcc	ttccctgttcc	cccagccatg	3480
agtgaatatg	ctgaatgagg	acctttgtaa	gttctgatg	agtagcatgt	taggagaatg	3540
aagcactaat	cccagagcta	atggaccttc	ctttcccttc	agtactgtaa	ggagttcttc	3600
tctcaatgtg	tgacctgccc	tgtctatcac	cagtgggtat	ctgagctt	ggcctctgt	3660
atgagatgtg	aaaagagaga	attgatgtcc	ctgactaaca	gcatcagacc	ttttgacacc	3720
tgcttgacca	ggatttggat	aaaaggagaa	tttctgcagg	aaaataactc	ttagaaaaga	3780
aacttaggaa	tacagagatt	tgacagagt	gctgatgtca	aggagaacaa	ggatgcagaa	3840
gaaactcaag	atgtatgtat	taaaacaaaa	gaacaataac	ctgaagggac	catgattctg	3900
ttatttgtata	taacacaagg	aaatgccccca	gattctc	taaaagat	aatgtacata	3960
ttaagtatac	tagcctttat	agttactgct	atctacatgt	ttatcaaaat	aaaagactat	4020
ttttttctaa	aaca					4034

<210> 25
<211> 4074
<212> DNA
<213> Homo sapien

<400> 25									
acactgggtt	cgagttccca	acctcagg	tc atctggccgt	ctcagcc	caaagt	gtcg	60		
ggattacagg	cgtgagccac	cgtgc	cgc	cgt	tat	ttttaaa	ggttag	ctca	120
cctaagactt	ccgcag	ctg	ggc	agtaac	aagat	aggca	tgat	gcac	180
ggattccaga	ccc	cctgt	tgcat	agttt	ccc	agg	ttt	gact	240
ctcattttt	tctggat	agg	tctac	ctgca	agtc	ggattt	ccc	agg	300
gagcggc	agc	aatac	c	tct	ttt	ggat	ttt	ggat	360
gtttggctca	ggc	gaat	gaa	att	gtc	tat	c	tca	420
atatgtttga	agtgg	caatc	atg	agg	ggg	atg	aa	gtc	480
ctata	ag	atc	ttc	atc	ttt	ttt	ttt	ttt	540
acaggac	ac	gtt	gtt	gtt	gtt	gtt	gtt	gtt	600
ccacagt	cat	tg	gaa	agg	gg	gg	gg	gg	660
gcatgtt	cct	ct	gtt	ttt	ttt	ttt	ttt	ttt	720

ccctggagag ctcccttcagc atcttcagct tctgctggga gaatcttgc aagaccaact	780
gaggaggaag gtggggcaga gaggagctc tcaggccccca ggggctcttc aggaggatc	840
cctagatttg ttccatcatcg tatcactaat ggaccagtat tctgtggcgt gggtggagca	900
gtagcaagac acgaatatac agttattggc caaaaagtga gtctgcggc cagaatgata	960
actgcttatac caggtttgggt gtcctgtatc gaggttaacat atctaagatc catgtcacct	1020
gcttacaact tcaagaaaact cccagagaaa atgatgaaaa acatctccaa cccagggaaag	1080
atatatgaat atcttggcca cagaagatgt ataatgtttg gaaaaagaca tttggcaaga	1140
aagagaaaaca aaaatcaccc tttgttagga gtgttaggtg ctccctgtct ctctacagac	1200
tgggagaaaag aatttggaaagc cttccaaatg gcacagcaag ggtgtttgca ccagaagaag	1260
ggacaagcag ttctgttatga aggtggaaaa ggctatggaa aaagccagct gttggctgaa	1320
ataaaacttgc tggcacagaa agaaggcat agctaccctt cacaggtgtt ttggaaaccc	1380
actttattgt gaggtcctat gccaggacct tctctctaag gacgtgttgc tcttcatgt	1440
cctacaaaag gaggaagagg aaaacagcaa gtggaaacc ctctcagcca atgccatgaa	1500
atccataatg tatagtattt ctcctgcca ctctgaggaa ggccaggaac tttatgtctg	1560
cacagtcaag gatgatgtga acttggatac agtacttctc ctacccttt tgaaagaaaat	1620
agcagtaagc caactggatc aactgagccc agaggaacag ttgctggtca agtgtgctgc	1680
aatcatttgt cactccttcc atatagattt gctgcagcac ctctgcctg gctggataa	1740
aaataagcta cttcaggctct tgagagctct tgtggatata catgtgctct gctggctga	1800
caagagccaa gagcttcctg ctgagcccat attaatgcct tccctatacg acatcattga	1860
tggaacccaa gagaagaaga caaaagttaga tggtggtca gcctcttttc tcaggctaca	1920
agaagaatta tccctaccac aaactgaggt gttggattt ggagtgcctc tgctacgggc	1980
agctgcttgg gagctctggc ccaaggaaca acagatagct ctgcacccctt aatgtgcctg	2040
ctttctccaa gttttggcct gccgctgtgg gagctgccat ggaggagact ttgtccccctt	2100
tcatcatttt gcagtttggtt ctactaagaa ttccaagggg acctctcgat tctgtactta	2160
cagagatact ggctcagtgc taacacaagt gatcacagaa aaattgcagc tgccttctcc	2220
ccaagaacag aggaagagtt cctagatcaa gtgaagagga agctggctca gaccagccct	2280
gagaaagacc tggtgaccac aaagccttgt cactgttaagg atatcctgaa gttagtgctc	2340
ttaccctca cccagcattt gttggtcgtt ggagaaacca cctgtgcatt ttattacctg	2400
ctggaggctg cggctgcctg cttggacctg tcagataatt atatggctcg tttcaacatg	2460
ggacgtatca cttagccaa aaaattggct agggaaagccc ttcgactgct gaaaaggaat	2520
ttcccttggc cctgggtttgg tgccttttc cagacattcc tggaaaagta ttggcattcc	2580

tgtaccctga	gccaacacctcc	aaacgaccct	agtgagaagt	tgtctctacc	tatgtggagc	2640
tctctcagtt	ctcccagagt	gtgggcatca	aggacaagtg	gctgcactgt	gagcagatgg	2700
ccattcagaa	aagcagtttta	tgttggttct	ccagggaggg	gttgttggcc	acagctcagc	2760
tcatgcaggc	cctggcctac	accaagctct	gccttggtca	tcttgacttc	tccatcaagc	2820
tgggattgct	gtgtcgcccc	tttagtgagt	gtctgcgttt	cgttcaagtc	tacgagcaca	2880
gccgtgttct	aacctctcag	agcaatgtca	tgctgggggt	ccactcctcc	ctggccatgt	2940
ggtaatgtct	tactcaaggg	ctgtggaaaa	ggatagacat	ttatgtcatt	taagctgtct	3000
ctccccacca	gacaggactt	ttgaacctct	ctaaccaact	tttaaagacc	attcacctcc	3060
cataccctcc	catcttatta	gaagggtctct	tgtcctttaa	caggtttgg	cctataggtc	3120
aagggttacg	tttagggtta	catttaactg	ctagagtaac	ccatagcaag	gctgaatata	3180
attggtctcc	ttttaagttt	ccttgtatgt	gagtttagtag	ccttggtcac	tttctagcat	3240
cacaattctg	attgtccatg	aggtcttaga	gccttaaaga	agtgtatgatt	ttaagcaaaa	3300
gtcatggtgg	gtaaggcagcg	gatattgctg	caagctgtta	ctctttcc	ccaggtttgc	3360
ccaggaatca	cagtgggacc	tgtttaagca	ctatttctcc	aacgcttgca	gttggtgaaa	3420
agaaccaatg	cctcgctatt	tgggcacat	ggctttgtcc	gattcctaga	atgcccattgt	3480
ttaatgttac	agaaaatgcc	agagggtatac	ttcatgcata	ttcctctaga	gcttcacagc	3540
caaaccctt	aggcttattt	tgccatcagt	aactccttcc	tgttccccc	gccatgagtg	3600
aatatgctga	atgaggacct	tttactgtaa	ggagttcttc	tctcaatgtg	tgacctgccc	3660
tgtctatcac	cagtgggtat	ctgagcttaa	ggcctctgta	atgagatgtg	aaaagagaga	3720
attgatgtcc	ctgactaaca	gcatcagacc	tttgacacc	tgcttgcacca	ggatttggat	3780
aaaaggagaa	tttctgcagg	aaaataactc	ttagaaaaga	aacttaggaa	tacagagatt	3840
tgacagagtg	gtgtatgtca	aggagaacaa	ggatgcagaa	gaaactcaag	atgtatgtat	3900
taaaacaaaa	gaacaataac	ctgaagggac	catgattctg	ttattgtata	taacacaagg	3960
aaatgccccca	gattctcctt	taaaagatat	aatgtacata	ttaagtatac	tagcctttat	4020
agttactgct	atctacatgt	ttatcaaaaat	aaaagactat	ttttttctaa	aaca	4074

<210> 26
<211> 3591
<212> DNA
<213> Homo sapien

<400> 26						
acactgggtt	cgagttccca	acctcaggtc	atctgcccgt	ctcagcctcc	caaagtgctg	60
ggattacagg	cgtgagccac	cgtgcctggc	cgtaaaggat	tatTTTaaa	ggtttagctca	120

cctaagactt ccgcagctga gggcagtaac aagataggca tgatgcacag agccatgtgg	180
ggattccaga cccctcctgt tgcatacgattt cccagttgaa ttgactctt ctccatttat	240
ctcattttt tctggatagg tctacactgca agtcggattt cccaggttat tggtggagat	300
gagcggcagc aataacctt ggtgattggg caggttgttag tgatgtccag ttagctcagc	360
gtttggctca ggcgaatgaa attgtccatat cctggaaactg ctggatgctt tgcaaggagt	420
atatgtttga agtggcaatc atgagggagg atgaagctgt gaagattgtt gaaggccagc	480
ctatagagta tgtatctgaa ttccgtacga tgactctggt ttgggtcagc ctagagttcc	540
acaggacagc gtggatgttg catttgcgtc atcttatcca ggaggctgcc ttatacatct	600
ccacagtcat tgagaaaggg ggcggccagc tgagtcggat ctttatgttt gagaaaggct	660
gcatgttccct ctgtgttttc ggccttcctg gtgataagaa gccagacgag tgtgcacatg	720
ccctggagag ctcccttcagc atcttcagct tctgctggga gaatcttgc aagaccaact	780
gaggaggaag gtggggcaga gaggagcttc tcaggccccca ggggctttc aggaggatc	840
cctagatttgc ttccatcag tatcactaat ggaccagtat tctgtggcgt ggttggagca	900
gtagcaagac acgaatatac agttattggc caaaaagtga gtcttgcggc cagaatgata	960
actgcttatac caggtttgggt gtcctgtat gaggttaacat atctaagatc catgctacct	1020
gcttacaact tcaagaaact cccagagaaa atgatgaaaa acatctccaa cccagggaaag	1080
atatatgaat atcttggcca cagaagatgt ataatgtttt gaaaaagaca tttggcaaga	1140
aagagaaaca aaaatcaccc tttgttagga gtgttaggtg ctccctgtct ctctacagac	1200
tgggagaaag aatttggaaagc cttccaaatg gcacagcaag ggtgtttgca ccagaagaag	1260
ggacaagcag ttctgtatga aggtggaaaa ggctatggaa aaagccagct gttggctgaa	1320
ataaaacttcc tggcacagaa agaaggcat agtaccctt cacaggtgtt ttggaaacccc	1380
actttattgt gaggtcctat gccaggacct tctctctaag gacgtgttgc tctttcatgt	1440
cctacaaaag gaggaagagg aaaacagcaa gtggaaacc ctctcagcca atgcccataatgaa	1500
atccataatg tatagttattt ctcctgccaa ctctgaggaa ggccagggaaac tttatgtctg	1560
cacagtcaag gatgtatgtga acttggatac agtacttctc ctacccttt tgaaagaaat	1620
agcagtaagc caactggatc aactgagccc agaggaacag ttgctggcgtca agtgtgtgc	1680
aatcattgggt cactccttcc atatagattt gctgcagcac ctccctgcctg gctggataa	1740
aaataagcta cttcaggatct tgagagctct tggatata catgtgtctt gctggatctga	1800
caagagccaa gagcttcctg ctgagccat attaatgcct tcctctatcg acatcattga	1860
tggAACAAA gagaagaaga caaagttaga tgggtggcgtca gcctctttc tcaggatcata	1920

agaagaatta	tccctaccac	aaactgaggt	gttggattt	ggagtgcctc	tgctacgggc	1980
agctgcttgg	gagctctggc	ccaaggAACa	acagatAGCT	ctgcacCCtG	aatgtgcctG	2040
ctttctccaa	gttttggcct	gccgctgtgg	gagctGCCat	ggaggagact	ttgtccccTT	2100
tcatcattt	gcagtttgtt	ctactaAGAA	ttCCAAGGGG	acctctcgat	tctgtactta	2160
cagagatact	ggctcagtgc	taacacaAGT	gatcacAGAA	aaattgcAGC	tgccttcTCC	2220
ccaagaACAG	aggaAGAGTT	cctagatCAA	gtGAAGAGGA	agctggctCA	gaccAGCCt	2280
gagaaAGACC	tgttgaccac	aaAGCCTTGT	caCTGTAGG	atATCCTGAA	gttagtgctc	2340
ttaccCCTCA	cccAGCATTG	cttggTCGTT	ggagaaACCA	cctgtgcatt	ttattacCTG	2400
ctggaggCTG	cggctgcCTG	cttggacCTG	tcAGATAATT	atATGGTCTG	tttcaACATG	2460
ggacgtatCA	ctttAGCCAA	aaaATTGGCT	aggAAAGCCC	ttcGACTGCT	gAAAAGGAAT	2520
ttccCCTTGA	cctggTTTGG	tgtcCTTTc	cAGACATTCC	tggAAAAGTA	ttggcATTCC	2580
tgtaccCTGA	gccaACCTCC	aaACGACCCt	agtGAGAAGT	tgtctctacc	tatgtggAGC	2640
tctctcAGTT	ctcccAGAGT	gtggGCATCA	aggACAAgTg	gctgcactgt	gagcAGATGG	2700
ccattcAGAA	aAGCAGTTA	tgttgTTCT	ccAGGGAGGG	gttGTTGGCC	acagctcAGC	2760
tcatgcAGGC	cctggCCTAC	accaAGCTCT	gccttggTC	tcttgactTC	tccatcaAGC	2820
tgggattGCT	gtgtCGGCC	tttagtgAGT	gtctgcgtt	cgttcaAGTC	tacgAGCACA	2880
gccgtgttCT	aacctCTCAg	agcaatgtCA	tgtGGGGGT	ccactCCtCC	ctggCCatGT	2940
ggtttGCCCA	ggaatCACAG	tggGACCTGT	ttaAGCACTA	tttctccAAc	gcttgcAGTT	3000
ggtgAAAAGA	accaATGCCT	cgttatTTGG	tgcACATGGC	tttGTCGAT	tcctAGAATG	3060
ccatgtgtTA	atgttACAGA	aaatGCCAGA	gggtatCTTC	atgcataTTc	ctctAGAGCT	3120
tcacAGCCAA	accCTTGAGT	actgtAAGGA	gttCTTCTCT	caatgtGTGA	cctGCCtGT	3180
ctatcAccAG	tggGTatCTG	agcttaAGGC	ctctgtAATG	agatgtGAAA	agAGAGAATT	3240
gatgtCCCTG	actAACAGCA	tcAGACCTT	tgACACCTG	ttgACCAGGA	tttggATAAA	3300
aggAGAATTt	ctgcAGGAAA	ataACTCTTA	gAAAAGAAC	ttagGAATAC	agAGATTG	3360
cagAGTGGCT	gatgtCAAGG	agaACAAAGGA	tgcAGAAGAA	actCAAGATG	tatgttATAA	3420
aacAAAAGAA	caataACCTG	aAGGGACCAT	gattCTGTtA	ttgttatATAA	cacaAGGAAA	3480
tgccccAGAT	tctCCTTAA	aAGATAATA	gtACATATTA	agtataACTAG	cctttatAGT	3540
tactgtatC	tacatgtTTA	tcaAAATAAA	agactATTT	tttctAAAC	a	3591

<210> 27
 <211> 5050
 <212> DNA
 <213> Homo sapien

<400> 27
 acactgggtt cgagttccca acctcaggc atctgcccgt ctcagccccc caaagtgc 60
 ggattacagg cgtgagccac cgtgcctggc cgtaaggtat tattttaaa ggttagctca 120
 cctaagactt ccgcagctga gggcagtaac aagataggca tgatgcacag agccatgtgg 180
 ggattccaga cccctccctgt tgcatagttt cccagttgaa tttgactctt ctccatatt 240
 ctcattttt tctggatagg tctacctgca agtccggattt cccaggttat tggtggagat 300
 gagcggcagc aataacctt ggtgattggg caggttgtag tgatgtccag tttagctcagc 360
 gtttggctca ggcgaatgaa attgtcctat cctggaaactg ctggatgctt tgcaaggcgt 420
 atatgttga agtggcaatc atgaggaggat atgaagctgt gaagattgtat gaaggccagc 480
 ctatagagta tgtatctgaa ttccgtacga tgactctggt tttggtcagc cttagattcc 540
 acaggacagc gtggatgttg catttggc atcttaccca ggaggctgcc ttatacatct 600
 ccacagtcat tgagaaaggg ggcggccagc tgagtcggat ctttatgtt gagaaaggct 660
 gcatgttcct ctgtgtttc ggccttcctg gtgataagaa gccagacgag tgtgcacatg 720
 ccctggagag ctcccttcagc atcttcagct tctgctgggaa gaatcttgct aagaccaact 780
 gaggaggaag gtggggcaga gaggagcttc tcaggccccaa ggggctttc aggccaggatc 840
 cctagatttg tttccatcag tatcactaat ggaccagtat tctgtggcgt ggttggagca 900
 gtagcaagac acgaatatac agttattggc ccaaagtga gtcttgcggc cagaatgata 960
 actgcttatac caggtttgggt gtccctgtat gaggtacat atctaaagatc catgtacatc 1020
 gcttacaact tcaagaaaact cccagagaaa atgatgaaaa acatctccaa cccagggaaag 1080
 atatatgaat atcttggcca cagaagatgt ataatgtttt gaaaaagaca tttggcaaga 1140
 aagagaaaaca aaaatcaccc tttgttagga gtgttaggtg ctccctgtct ctctacagac 1200
 tgggagaaag aatttggaaagc cttccaaatg gcacagcaag ggtgtttgca ccagaagaag 1260
 ggacaagcag ttctgtatga aggtggaaaa ggctatggaa aaagccagct gttggctgaa 1320
 ataaactttc tggcacagaa agaaggccat agctaccctt cacaggtgct ttggaaaccc 1380
 actttattgt gaggtccat gccaggacat tctctctaag gacgtgttc tctttcatgt 1440
 cctacaaaag gaggaagagg aaaacagcaa gtggaaacc ctctcagcca atgccatgaa 1500
 atccataatg tatagtattt ctccctggccaa ctctgaggaa ggccaggaac tttatgtctg 1560
 cacagtcaag gatgtatgtca acttggatac agtacttctc ctacccttt tgaaagaaat 1620
 agcagtaagc caactggatc aactgagccc agaggaacag ttgctggcgtca agtgtgtgc 1680
 aatcatgggt cactccttcc atatagattt gctgcagcac ctccctgcctg gctggataa 1740
 aaataagcta cttcaggctt tgagagctct tggatata catgtgtctt gctggctgaa 1800

caagagccaa gagcttcctg ctgagccat attaatgcct tcctctatcg acatcattga 1860
 tggAACAAA gagaagaaga caaagttaga tggtgggtca gcctctttc tcaggctaca 1920
 agaagaatta tccctaccac aaactgaggt gttggattt ggagtgcctc tgctacgggc 1980
 agctgcttgg gagctctggc ccaaggaaca acagatagtctgcacccat 2040
 ctttctccaa gttttggcct gccgctgtgg gagctgccat ggaggagact ttgtccccctt 2100
 tcatcatttt gcagtttgtt ctactaagaa ttccaagggg acctctcgat tctgtactta 2160
 cagagatact ggctcagtgc taacacaagt gatcacagaa aaattgcagc tgccttctcc 2220
 ccaagaacag aggaagagtt cctagatcaa gtgaagagga agctggctca gaccagccct 2280
 gagaaagacc tggtgaccac aaagccitgt cactgtaagg atatcctgaa gtttagtgctc 2340
 ttaccctca cccagcattg cttggtcgtt ggagaaacca cctgtgcatt ttattacctg 2400
 ctggaggctg cggctgcctg cttggacctg tcagataatt atatggctg tttcaacatg 2460
 ggacgtatca cttagccaa aaaattggct agaaaagccc ttgcactgt gaaaaggaat 2520
 ttcccttggc cctggtttgg tgccttttc cagacattcc tggaaaagta ttggcattcc 2580
 tgtaccctga gccaacctcc aaacgaccct agtgagaagt tgtctctacc tatgtggagc 2640
 tctctcagtt ctccccagagt gtgggcatca aggacaagt gctgcactgt gagcagatgg 2700
 ccattcagaa aagcagtttgc tgggggttct ccagggaggg gttgtggcc acagctcagc 2760
 tcatgcagggc cctggcctac accaagctct gcctggctca tcttgacttc tccatcaagc 2820
 tgggattgttgc gtgtcgcccc ttttagtgagt gtctgcgttt cggtcaagtc tacgagcaca 2880
 gccgtgttct aacctctcag agcaatgtca tgctgggggt ccactcctcc ctggccatgt 2940
 ggtaatgtct tactcaaggg ctgtggaaaa ggatagacat ttatgtcatt taagctgtct 3000
 ctccccacca gacaggactt ttgaacctct ctaaccaact tttaaagacc attcacctcc 3060
 cataccctcc catcttatta gaagggtctc tgcctttaa caggtttgg cctataaggc 3120
 aagggttacg ttttagggtaa catttaactg cttagttaac ccatagcaag gctgaatata 3180
 attggctcc ttttaagttt cttgttatgt gagtttagtag cttggctcac tttctagcat 3240
 cacaattctg attgtccatg aggtctttaga gccttaaaga agtgtatgatt ttaagcaaaa 3300
 gtcatggtgg gtaaggcagcg gatattgctg caagctgtta ctctttccct ccaggtttgc 3360
 ccaggaatca cagtgggacc tggtaagca ctatttctcc aacgcttgca gttggtaaaa 3420
 agaaccaatg cctcgctatt tggtgacat ggctttgtcc gattcctaga atgcccattgt 3480
 ttaatgttac agaaaatgcc agagggtatc ttcatgcata ttccctctaga gcttcacagc 3540
 caaaccccttg aggtacctgt ttctcagctg tccttgact aacacctgtatc tcaacttagtt 3600

40

ctaccctatg	gtgcttttc	taccacctgc	atctcttccct	ttttccctt	ttactggctc	3660
tgtttccctt	tactcttga	atccttggtt	tctccaccta	gaaagttct	acctacaccta	3720
tgtatccctc	ccgatattat	tgcatactagt	tctggactgg	gtttcttaac	tttccacctt	3780
tgccagctgc	tacccagtagt	cattaaaata	ttaacatcca	gccttgctca	atggacctgt	3840
agtctatggt	tcagtcata	atttgataca	gctccctcca	gcccttctga	gtctaaaaca	3900
cattccaatt	cctctgtttt	ccaggottat	tttgcacatca	gtaactcctt	cctgtcccc	3960
cagccatgag	tgaatatgct	gaatgaggac	cttgcatagt	tctgatgaag	tagcatgtta	4020
ggagaatgaa	gcactaatcc	cagagctaat	ggaccttcct	ttcccttcag	tactgtaagg	4080
agttcttctc	tcaatgtgtg	acctgcctg	tctatcacca	gtgggtatct	gagottaagg	4140
cctctgtaat	gagatgtgaa	aagagagaat	tgtatgcct	gactaacagc	atcagacctt	4200
ttgacacctg	cttgaccagg	atttggataa	aaggagaatt	tctgcaggaa	aataactctt	4260
agaaaaagaaa	cttaggaata	cagagtaagc	atttcttcct	ggaagccttg	tgtgagagac	4320
ataaaagacag	tctcagattc	ttactcacaa	gcagtcaaag	gctgcaccc	tgaaataaaa	4380
agggacacac	agatgtaagg	agtttgtcct	tgctccagag	gtaagtataa	tcctcttcct	4440
agtgcgtggc	cctgcctgga	cagataggaa	tcccttctat	tgttaaacag	caatttcttc	4500
agcttctctc	agctctttgt	ttcagtattg	gtaactcttt	ggcatagaaaa	gttcttcctt	4560
gcttttagcc	aaagcagttg	ggttgttcc	ttgaagtaac	tggatggtca	ctaaggagag	4620
aaaaaggtct	tagaagtcac	aatgtaatgt	ctatgaaggt	gaatgataag	attaggcaag	4680
aaaaggagag	gaaagaatat	agttcccttc	ctcagaggcc	tgcaaatctt	ctttccatg	4740
gctgctatTTT	aactttgtaa	ttgctgagga	cattcttgc	atttgtgaca	ttctttgtgt	4800
tccttcttc	aggatttgac	agagtggctg	atgtcaagga	gaacaaggat	gcagaagaaa	4860
ctcaagatgt	atgtattaaa	acaAAAgaac	aataacctga	agggaccatg	attctgttat	4920
tgtatataac	acaaggaaat	gccccagatt	ctcccttaaa	agatataatg	tacatattaa	4980
gtatactagc	ctttatagtt	actgctatct	acatgtttat	caaaataaaa	gactatTTT	5040
ttctaaaaca						5050

<210> 28
<211> 4658
<212> DNA
<213> Homo sapien

<400> 28
acactgggtt cgagttccca acctcaggc acctgcccgt ctcagcctcc caaagtgcgtg 60
ggattacagg cgtgagccac cgtgcctggc cgtaaggtat tattttaaa gtttagctca 120

41

cctaagactt	ccgcagctga	ggcagtaac	aagataggca	tcatgcacag	agccatgtgg	180
ggattccaga	cccctcctgt	tgcatagttt	cccagttgaa	tttgactctt	ctccatttat	240
ctcattttt	tctggatagg	tctacctgca	agtccggattt	cccaggttat	tgtggagat	300
gagcggcagc	aataacctt	ggtgattggg	caggttgtag	tcatgtccag	ttagctcagc	360
gttggctca	ggcgaatgaa	attgtcttat	cctggaaactg	ctggatgttt	tgcaaggcagt	420
atatgttga	agtggcaatc	atgagggagg	atgaagctgt	gaagattgtat	gaaggccagc	480
ctatagagta	tgtatctgaa	ttccgtacga	tgactctgg	tttggtcagc	ctagagttcc	540
acaggacagc	gtggatgttg	cattttgttc	atcttatcca	ggaggctgcc	ttatacatct	600
ccacagtcat	tgagaaaaggg	ggcgccagc	tgagtcggat	ctttatgttt	gagaaaaggct	660
gcatgttcct	ctgtgtttc	ggccttcctg	gtgataagaa	gccagacgag	tgtgcacatg	720
ccctggagag	ctccttcagc	atcttcagct	tctgctggg	aatcttgc	aagaccaact	780
gaggaggaag	gtggggcaga	gaggagcttc	tcaggcccc	ggggctcttc	aggcaggatc	840
cctagattt	tttccatca	tatcactaat	ggaccagtat	tctgtggcgt	ggttggagca	900
gtagcaagac	acgaatatac	agttattggc	ccaaaaagtga	gtcttgcggc	cagaatgata	960
actgcttatac	caggtttgg	gtcctgtat	gaggttaacat	atctaagatc	catgctacat	1020
gcttacaact	tcaagaaact	cccagagaaa	atgatgaaaa	acatctccaa	cccagggaaag	1080
atatatgaat	atcttggcca	cagaagatgt	ataatgttt	aaaaaagaca	tttggcaaga	1140
aagagaaaaca	aaaatcaccc	tttggtagga	gtgttaggt	ctccctgtct	ctctacagac	1200
tgggagaaag	aatttggaaagc	cttccaaatg	gcacagcaag	ggtgtttgca	ccagaagaag	1260
ggacaagcag	ttctgtatga	agggtggaaa	ggctatggaa	aaagccagct	gttggctgaa	1320
ataaaactt	tggcacagaa	agaaggcat	agctaccctt	cacaggtgt	ttggaaaccc	1380
actttatgt	gaggtcctat	gccaggacct	tctctctaag	gacgtgtgc	tcttcatgt	1440
cctacaaaag	gaggaagagg	aaaacagcaa	gtggaaacc	ctctcagcca	atgcctatgaa	1500
atccataatg	tatagtattt	ctcctgcca	ctctgaggaa	ggccaggaac	tttatgtctg	1560
cacagtcaag	gatgtgtga	acttggatac	agtacttctc	ctacccttt	tgaaagaaat	1620
agcagtaagc	caactggatc	aactgagccc	agaggaacag	ttgctggtca	agtgtgtgc	1680
aatcatttgtt	cactccttcc	atatagtattt	gctgcagcac	ctcctgcctg	gctggataa	1740
aaataagcta	cttcaggatct	tgagagctct	tgtggatata	catgtgtct	gctggatctga	1800
caagagccaa	gagtttcctg	ctgagccat	attaatgcct	tcctctatcg	acatcattga	1860
tggAACAAA	gagaagaaga	caaagttaga	tgggggtca	gcctcttcc	tcaggctaca	1920
agaagaatta	tccctaccac	aaactgaggt	gttggatatt	ggagtgcctc	tgctacgggc	1980

agctgcttgg gagctctggc ccaaggaaca acagatacg ctgcacccgg aatgtgcctg	2040
ctttcctccaa gttttggcct gccgctgtgg gagctgccat ggaggagact ttgtccccctt	2100
tcatcatttt gcagtttgtt ctactaagaa ttccaagggg acctctcgat tctgtactta	2160
cagagatact ggctcagtgc taacacaagt gatcacagaa aaattgcagc tgccttctcc	2220
ccaagaacag aggaagagtt cctagatcaa gtgaagagga agctggctca gaccagccct	2280
gagaaagacc tggtgaccac aaagccttgt cactgtaagg atatcctgaa gtttagtgctc	2340
ttaccctca cccagcattg cttggcgtt ggagaaacca cctgtgcatt ttattacctg	2400
ctggaggctg cggctgcctg cttggacctg tcagataatt atatggtctg tttcaacatg	2460
ggacgtatca ctttagccaa aaaattggct agggaaagccc ttgcactgt gaaaaggaat	2520
ttcccttggc cctgggttgg tgccttttc cagacattcc tggaaaagta ttggcattcc	2580
tgtaccctga gccaacctcc aaacgaccct agtgagaagt tgtctctacc tatgtggagc	2640
tctctcagtt ctcccagagt gtgggcatca aggacaagtg gctgcactgt gagcagatgg	2700
ccattcagaa aagcagttt ttttttttcc ccagggaggg gttgttggcc acagctcagc	2760
tcatgcaggc cctggcctac accaagctct gccttggtca tcttgacttc tccatcaagc	2820
tgggattgct gtgtcgcccc tttagtgagt gtctgcgtt cgttcaagtc tacgagcaca	2880
gccgtttct aacctctcag agcaatgtca tgctgggggt ccactcctcc ctggccatgt	2940
ggtaatgtct tactcaaggg ctgtggaaaa ggatagacat ttatgtcatt taagctgtct	3000
ctccccacca gacaggactt ttgaacctct ctaaccaact tttaaagacc attcacctcc	3060
cataccctcc catcttatta gaagggtct tgcctttaa caggtttgg cctataggtc	3120
aagggttacg tttagggta catttaactg ctagagtaac ccatagcaag gctgaatata	3180
attggctcc tttaagttt ctttgtatgt gagtttagtag cttggtcac tttcttagcat	3240
cacaattctg attgtccatg aggtcttaga gccttaaaga agtgtatgatt ttaagcaaaa	3300
gtcatggtgg gtaagcagcg gatattgctg caagctgtt ctctttcct ccaggtttgc	3360
ccaggaatca cagtgggacc tgtttaagca ctatttctcc aacgcttgca gttggtgaaa	3420
agaaccaatg cctcgctatt tggtgacat ggctttgtcc gattcctaga atgccatgtg	3480
ttaatgttac agaaaatgcc agagggatgc ttcatgcata ttccctctaga gcttcacagc	3540
caaacccttg aggtacctgt ttctcagctg tcctttgact aacacctgat tcacttagtt	3600
ctaccctatg gtgctcttcc taccacctgc atctcttcct tttttccctt ttactggctc	3660
tgtttccctt tactcttga atcctttgtt tctccaccta gaaagttctt acctaccttta	3720
tgtatcccttc ccgatattat tgcacatgt tctggactgg gtttcttaac ttccacctt	3780

43

tgccagctgc tacccagttat cattaaaata ttaacattta gccttgcgtca atggacctgt	3840
agtctatggt tcagtctata atttgataca gctccctcca gcccttctga gtctaaaaca	3900
cattccaatt cctctgtttt ccaggcttat tttgccatca gtaactcctt cctgttcccc	3960
cagccatgag tgaatatgct gaatgaggac ctggtaagt tctgtatgaa tagcatgtta	4020
ggagaatgaa gcactaatcc cagagctaat ggacccctt ttcccccag tactgttaagg	4080
agttttctc tcaatgtgtg acctgcctg tctatcacca gtgggtatct gagcttaagg	4140
cctctgtat gagatgtgaa aagagagaat tgatgtccct gactaacagc atcagacctt	4200
ttgacacctg cttgaccagg atttggataa aaggagaatt tctgcaggaa aataactctt	4260
agaaaaagaaa ctttaggaata cagagtaagc atttcttcctt ggaagcctt gttgagagac	4320
ataaaagacag tctcagattc ttactcacaa gcagtcaaag gctgcaccc tcgaaataaaa	4380
agggacacac agatgttaagg agtttgttct tgctccagag gatttgacag agtggctgat	4440
gtcaaggaga acaaggatgc agaagaaact caagatgtat gtattaaaac aaaagaacaa	4500
taacctgaag ggaccatgat tctgttattt tatataacac aaggaaatgc cccagattct	4560
cctttaaaag atataatgta catattaagt atactagcct ttatagttac tgctatctac	4620
atgtttatca aaataaaaaga ctatttttt ctaaaaca	4658

<210> 29
<211> 1920
<212> DNA
<213> Homo sapien

<400> 29	
ctccctccctc ctccactctg ctcaggccc tctactctt tttttttttt aaccgctacg	60
ccacagtccc cgggagaatt cagatccaa cggggcttc cggttctgt agtggcttt	120
gcctgtgtct ggtctgagga cggccggaaag gcattgcact gaggctaagg gaaaggtctc	180
tggagggagc ctcaggaaga gcaaatggag gccagagact ggcaggagcg cgccagcgca	240
ggatttaatc ccgacgagcg gattcagagc cgtgcttata taaagcttca ggaagegccc	300
ttccgacgat gaggtcgaca cgcgagaggc gacctcaaga gccccggcgc cagggatctg	360
tgcgccaagg gaggacggga gggagcaggc tcggcataat tccctggctcc aggctctgtt	420
ttgttggacc gagccactgt atttagctc acacaggaga attctggccc tggaaaatt	480
ggtctcagca tgctgccaag ctttctcatg gacgtcagcg aatcccaaca cactgtcggt	540
caaagccgtg ctggaagaaa caaaacagtt ctcctcggt agaactgaga ggggattgggt	600
ccagggcccc cggccatacc aaaatccagg ttgctcaagt ctctcataga aagtggcgta	660
gtatggcac ataactatgc acatcctccc gtgtacttta aatagtctct aaattacttc	720

gtaacaccta atccagtgt aatgcatagt aagtaattgt tatactgttt ttatTTTAC	780
tatTTTTGT tgtactTTTT tttaaaaaAG aaattcATTt gttaatATTt ttcggTCTTG	840
gggaACCCGC gtATATGGAG ggcctgCTAC atAGAGAAGA ctGAGGGATA ttCTGTGCA	900
ccgTTTCTAC ggATCCTCTA aATCGGCCTT tgTTTCAGC caggATTAG tgCCCAGCTG	960
tgtcTTTGG aggccccaca tggagCTAGC aaAGTTTGT aaATCGGGTT ttGCAAGAGG	1020
actgtCTGCT ccataCTGGG agtagTTACC gcaAAACTGCC CTATGAAATT ggTTGGGTT	1080
cTTACTGTtA gcatgtttAT tacTTTATCA gggCTCTCTG taggAGAGTC tatgAGAAA	1140
tcttCTGGTT tctgCTGAAA gaATCGTGTt ttGTTGGGT tttttcccG aaaaATATTa	1200
tTTTTAAAAA ctTTCTGTG ccCTGTTAA tCTCTCCCTT ggATCCACCT tCTGTGTGCT	1260
cataAAATCGT aaATCTGTAT tcAGACTTCT ggACTCGAGA cacGTAgATC cacCTGGTGG	1320
ttCTTCAGTC attTTAAGCC caAAACTCAA aATCTCCGA aATCAAATG ttTTAAACTTA	1380
taATCTCCAG ggtGTGACTC acGGGGGATG agGGGAGCAA ttCTCTCCCTT cccCGATAA	1440
agCTGGTTCT cCTGTCTGCT cATTGAAACGG ttCCACTGCG catCACAGCA tCTACATGCC	1500
taAAACCAACA cccCAGCATT ggCAACAGAT ATCTTCCTCT ccCTTGGCTG ctTCAGGACA	1560
ggGAAGAAAC atGCTTGCCC ttTTCTGACT CTtTAGTAAC tCTGGCCGAA tCTATCACAT	1620
tATTTACAT ctCTTTACAT ctTACTACTC cccCATCTTG gCTGTGTGTT ccCTACTGGC	1680
agtGATTTT GTTTATTCAt ttttgtaaaAC tgACACTTAG ttCAGTGTCC aATATAAGCT	1740
caACAATAGT ttATAAAGGA aaAGTTCTG cCTTGTGATTG ctTTAAACA ctATTAGAAA	1800
agACATAACC aaATTGCAAC atGATAAAAC aAccGCAAAC aAGGCTGAGA gaAGTGGTGA	1860
tttCTGGTGT cAGAGGGCAC aggACCTGG gcAGAAATCAG agATACGGTG tCTGTGCAgT	1920

<210> 30
<211> 6398
<212> DNA
<213> Homo sapien

<400> 30	
gcCTTCCCA agtGCTTGT aatGAATAGA aATGGAAACC aaaaaAAACG tataCAGGCC	60
ttcAGAAATA gtaATTGCTA ctATTTGTT ttcATTAGC cataGTTCTG gCTATAATT	120
tatCAAACTC accAGCTATA ttCTACAGTG aaAGCAGGAT tCTAGAAAGT cTCACTGTT	180
tATTTATGTC accATGTGCT atGATATATT tGGTTGAATT cATTGAAAT tagGGCTGGA	240
agtATTCAAG taATTCTTC tgCTGAAAAA atACAGTGTt ttGAGTTAG ggcCTGTTT	300
atCAAAGTTC taaAGAGCCT atCACTCTTC cATTGAGAC ATTtAAAT aATGACACTG	360
atTTAAACAT ttTTAAGTGT ctTTTGTAA cAGAGAGCCT gACTAGAAACA cAGCCCCCTCC	420

45

aaaaacccat gctcaaatta ttttactat ggcagcaatt ccacaaaagg gaacaatggg	480
tttagaaatt acaatgaagt catcaaccca aaaaacatcc cstatccctaa gaaggttatg	540
atataaaatg cccacaagaa atctatgtct gcttaatct gtctttatt gcttggaaag	600
gatggctatt acattttag ttttgctgt gaatacctga gcagtttc tcatccatac	660
ttatcctca cacatcagaa gtcaggatag aatatgaatc atttaaaaa ctttacaac	720
tccagagcca tgtgcataag aagcattcaa aacttgccaa aacatacatt tttttcaaa	780
tttaaagata ctctatTTT gtattcaata gctcaacaac tgtggcccc actgataaag	840
tgaagtggac aaggagacaa gtaatggcat aagttgttt ttcccAAgt atgcctgttc	900
aatagccatt ggatgtggga aatttctaca tctctaaaa ttacagaa aatacatagc	960
cagatagtct agcaaaagt caccaagtcc taaattgctt atccttactt cactaagtca	1020
tgaaatcatt ttaatgaaaaa gaacatcacc taggtttgtt ggTTTCTTT tttcttattc	1080
atggctgagt gaaaacaaca atctctgtt ctccttagca tctgtggact atttaatgt	1140
ccattattcc acactctatg gtccttacta aatacaaaat tgaacaaaaa gcagtaaaac	1200
aactgactct tcacccatat tataaaatata aatccaagcc agattagtca acatccataa	1260
gatgaatcca agctgaactg ggcctagatt attgagttca ggTTGGatca catccctatt	1320
tattaataaa cttagggaaag aaggccttac agaccatcag ttagctggag ctaatagaac	1380
ctacacttct aaagttcggc ctagaatcaa tgtggctta aaagctgaaa agaagcagga	1440
aagaacagtt ttcttcaata atttgcac CCTGTCactg gagaaaattt aagaatttgg	1500
gggtgttgtt agtaagttaa acacagcagc tgTTcatggc agaaatttatt caatacatac	1560
cttctctgaa tatcctataa ccaaagcaaa gaaaaacacc aaggggTTTg ttctcctcct	1620
tggagttgac ctcattccaa ggcagagctc aggtcacagg cacagggct gcGCCCAAGC	1680
ttgtccgcag ctttatgcag ctgtggagtc tggaaagactg ttgcaggact gctggcctag	1740
tcccagaatg tcagcctcat ttgcattta ctggctttt tgctgtatg tcatgctgac	1800
cttattgtta aacacaggtt tgTTTGTttt ttTCCACTC atggagacat gggagaggca	1860
ttatTTTaa gctggTTgaa agCTTAACC gataaAGcat tttagagaa atgtgaatca	1920
ggcagctaag aaagcataact ctgtccatta cgtaaAGaa aatgcacaga ttattaACTC	1980
tgcagtgtgg cattagtgtc ctggtaata ttccggataga tatgaataaa atatTTAAAT	2040
ggTATTGTTAAtagTTTCA ggacatATGC tataGCTTAT ttTTTATTATC ttTTGAAATT	2100
gctcttaata catcaaATCC tgatgtattc aatttATCAG atataAAATTt ttctAAATGA	2160
agcccAGTTA aatgtTTTG ttTTGTCAgT tataGTTAA gtttCTGATC ttTTGTCTA	2220
tgacgtttac taatctgcat ttTACTGTT atgaatttatt ttagacagca gtggTTCAA	2280

gcttttgcc actaaaaata ccttttattt tctcctcccc cagaaaagtc tataacctga	2340
agtatctatc caccaaactg tacttctatt aagaaatagt tattgtgtt tcttaatgtt	2400
ttgttattca aagacatatac aatgaaagct gctgagcagc atgaataaca attatatcca	2460
cacagatttgc atatattttg tgca gccc ttactt gatagt ataaaatgtc attgttttt	2520
aaataatagt tagtcaatgg acttctatca tagcttcct aaacttaggtt aagatccaga	2580
gctttgggt cataatatac tacatacaat taagttatct ttttctaagg gctttaaaat	2640
tcatgagaat aaccaaaaaaaaa ggtatgtgga gagttaatac aaacatacca tattcttgg	2700
gaaacagaga tgtggctctg cttgttctcc ataaggtaga aataacttcc agaatttgc	2760
taaacttagta agccctgaat ttgctatgtat tagggatagg aagagatttt cacatggcag	2820
actttagaat tcttcacttt agccagtaaa gtatctcctt ttgatcttag tattctgtgt	2880
attttaactt ttctgagttt tgcatgttta taagaaaaat cagcacaaag ggtttaagtt	2940
aaagcccttt tactgaaatt tgaaagaaac agaagaaaaat atcaaagtcc tttgtatttt	3000
gagaggatta aatatgattt acaaaagtta catggagggc tctctaaaac attaaattaa	3060
ttatTTTTG ttgaaaagtc ttactttagg catcattttt ttcctcagca actagctgtg	3120
aagcctttac tggctgttat gccagtcact ctgctagatt gtggagat ccagtgtcc	3180
cgtcttctcc gagcttagag ttggatgggg aataaagaca ggtaaacaga tagctacaat	3240
attgtactgt gaatgcttat gctggaggaa gtacaggaa ctattgggc acctaagagg	3300
agcacccacc ttgaattttg gggtagcag aggcattctg aaaaaagtca aagctaagcc	3360
acaatctata agcagtttag gaatttagcag aacgtgcgtg gtgaggagat gccaaggca	3420
agaagagaag agtattccaa acaggaggaa ttccaaagag agaagagtat cccaaacaac	3480
atttgcacaa acctgatggg gagagagaat gtgggggtggg gatggatgat gagactgaag	3540
aagaaagcca ggtctagata atcagtggcc ttgtacacca tggtaaagag tggtagacttg	3600
attctgttgtt aaacaggaaa gcagcacaat tcataatgat attttagaag actcccaactg	3660
gaatatggag aataaagtgg gagatgacta atcctggaaag cagggagaac atttttgagg	3720
aagttgcact attttgggtga aaatgtatgtt cataaacatg aagaattgtt ggtgtatgt	3780
acccctcttc taatTTTCCA gaagggtttt ggaagatata acataggaac attgacagga	3840
ctgacgaaag gagatgaaat acaccatata aattgtcaaa cacaaggcca gatgtctaat	3900
tatTTTGTt atgtgttgaa attacaattttt tttcatcagg aaacccaaaaa ctacaaaact	3960
tagTTTCCC aagtcccaga attctatctg tccaaacaat ctgtaccact ccacccat	4020
ccctacccccc gcatgtctgtt ccaaccccaa agtccaggc tatacacacg ggtaagacta	4080

gagcagttca agtttcagaa aatgagaaaag aggaactgag ttgtgctgaa cccataaaaa	4140
ataaacacat tctttgtata gattcttggaa acctcgagag gaattcacct aactcatagg	4200
tatttgatgg tatgaatcca tggctggct cggttttaa aaagccttat ctgggattcc	4260
ttctatggaa ccaagttcca tcaaagccc tttaaaagcc tacattaaaa acaaaattct	4320
tgctgcattg tatacaaata atgatgtcat gatcaaataa tcagatgcc tcatacaagt	4380
gaattacaaa atggataacc cactccaaaa aaaaaaaagct aaattctcag tagaacattg	4440
tgacttcattg agccctccac agccttggag ctgaggaggg agcactggtg agcagtaggt	4500
tgaagagaaa acttggcgct taataatcta tccatgtttt ttcatctaaa agagccttct	4560
ttttggatta ctttattcaa tttccatcaa ggaaattgtt agttccacta accagacagc	4620
agctgggaag gcagaagctt actgtatgta catggtagct gtgggaagga ggtttcttc	4680
tccaggtcct cactggccat acaccagtcc cttgttagtt atgcctggc atagaccccc	4740
gttgctatca tctcatattt aagtctttgg cttgtgaatt tatctattct ttcagcttca	4800
gcactgcaga gtgctggac tttgctaact tccatttctt gctggcttag cacattcctc	4860
ataggcccac ctctttctc atctggccct gctgtggagt caccttgccc cttcaggaga	4920
gccatggctt accactgcct gctaaggcctc cactcagctg ccaccacact aaatccaagc	4980
ttctctaaga tggcagac ttacaggca agcataaaaag gcttgatctt cctggacttc	5040
cctttacttg tctgaatctc acctccttca actttcagtc tcagaatgta ggcatttgc	5100
ctctttgccc tacatcttcc ttcttctgaa tcatgaaagc ctctcacttc ctcttgctat	5160
gtgctggagg cttctgtcag gttttagaat gagttctcat cttagcctag tagctttga	5220
tgcttaagtc cacctttaa ggataccctt gagattaga ccatgtttt cgcttgagaa	5280
agccctaatac tccagacttg ctttctgtg gatttcaaag accaactgag gaagtcaaaa	5340
gctgaatgtt gactttctt gaacatttcc gctataacaa ttccaattct cctcagagca	5400
atatgcctgc ctccaaactga ccaggagaaa ggtccagtgc caaagagaaa aacacaaaga	5460
ttaattttt cagttgagca catacttca aagtggtttgg tattttacttata tgaggtttc	5520
tgtcaagagg gtgagactct tcatactatcc atgtgtgcct gacagttctc ctggcactgg	5580
ctggtaacag atgcaaaaact gtaaaaatta agtgatcatg tattttaaacg atatcatcac	5640
atacttattt tctatgtaat gttttaattt tcccttaaca tactttgact gttttgcaca	5700
tggttagatat tcacatTTT ttgtgttggaa gttgatgcaa tcttcaaagt tatctacccc	5760
gttgcttattt agtaaaacta gtgttaatac ttggcaagag atgcagggaa tctttctcat	5820
gactcagcc ctatTTTttt attaatgcta ctaccctatt ttgagtaagt agtaggtccc	5880
taagtacatt gtccagagtt atactttaa agatatttag ccccatatac ttcttgaatc	5940

taaagtctata caccttgctc ctcatttctg agtgggaaag acatggaga gttatgttgac 6000
 aattttctg aagggttttg ccaagaaggt gaaaactgtcc tttcatctgt gtatgcctgg 6060
 ggctgggtcc ctggcagtga tgggggtgaca atgcaaagct gtaaaaacta ggtgcttagtg 6120
 ggcacctaattt atcatcatca tataacttatt ttcaagctaa tatgcaaaat cccatctctg 6180
 tttttaact aagtgttagat ttcaagagaaa atattttgt gttcacataa gaaaacagtc 6240
 tactcagctt gacaagtgtt ttatgttaaa ttggctggtg gtttggaaatg aatcatcttc 6300
 acataatgtt ttctttaaaa atattgtgaa tttaactcta attcttgttta ttctgtgtga 6360
 taataaagaa taaactaatt tctatatctc tctttatt 6398

<210> 31
 <211> 1314
 <212> DNA
 <213> Homo sapien

<400> 31
 aggtgcgggc gcccagccca gggcaggcgg gcagggctga gggcgccggat ccccaaccag 60
 gccccgcgca ctttcattgac gtttcagaac tgctccgagg caaaactcaga caactctctg 120
 aggacaacgt ccgcggccgc ggcgcccgc tctttcggg gccagggacc ggggtgtcgg 180
 tcctattcga aaggggacgga gaactacatt tcccgcatg ccatcgccca ctccgggcct 240
 ggcacggaaa gagctttcg cagccgaacg tcatttcgc tgctactg ggaccacggt 300
 ctgttagtcgt gagcggaggc ctggtatggc gcccggtttc cggttccgg cgacggaaat 360
 gacgctatca cggcgccca aggcgtcagt cgaggagtca aggcagcaat gaatcggtc 420
 ttgtgtgccc cggcgcccg ggcgtccgg ggcgtgaggc tcataggctg ggcttcccga 480
 agccttcatac cggtggccgg ttccggat cggggccacc ctggccgcga ggaaggaggac 540
 gaccctgacc gccccattga gtttcctcc agcaaagcca accctcaccg ctggtcggtg 600
 ggccatacca tggaaaggc acatcagcgg ccctggtgga aggtgctgcc cctcagctgc 660
 ttcctcgatcg ctggatcgac ctgaggaggag agagcgaggc ggaccaggatgg 720
 ttgagacagg tgtggggaga ggtgccagag ccoagtgtac gttctgagga gcctgagact 780
 ccagctgcct acagagcgag aacttgacgg ggtgcggct gggctggca ggaaggaggc 840
 cgacagccgc ctttcggatt tggatgtcact tttggccgtg actgtccctgg ctatcggtc 900
 gtcctcagca ctgaaggact tggatgtcact atggggact tggatgtcact gattcggtc 960
 aaggccggc agaatctcag cagatcgaa actgtccctc gcctggctct tggatgtccaa 1020
 ggattccatc ggcaagactt ctcagatcct tggggaaaggc ttcagttgca ctgtatgtc 1080
 ttggatgtc caagtcttttgc tataacataa tcatgtttcc aaagcacttc tggtgacact 1140

tgtcatccag tggtagttt caggtaattt gctttctgag atagaatatac tggcagaagt	1200
gtgaaactgt attgcattgt gcggcctgtg caaggaacac ttccacatgt gagttttaca	1260
caacaacaaa tgaaaataaaa tttaatttt ataatatggg attagatgtat tccc	1314

<210> 32
<211> 1124
<212> DNA
<213> Homo sapien

<400> 32	
tttcctcgct gcagtcattcc aatagccaag atacacggct aggtgatttgcagcgggag	60
ttaggtgtcc tcttggcgcc tgaccagat cggaaatttgcgttccatgttttggatgtccc	120
ttccccgagt tgccccccga ggtatgcggg gtcactcgct gctcgatgtt ccctccgaag	180
ggtcggacaa ggctccggag ccctgttagct gccctcccta ggagccccgg gtcttcactg	240
gccgaggtgc ccaccccgca gcattctggg agtggtagtt ttcttccttc agttcatcc	300
ctggctggcc agtgcggcaag actggcgaga ctacgattcc cagacgcggc agcgagtcgc	360
cggcacgtg gccgcaagga cgctggcccg gtgggcgggg gccggcaggt gctccgcagc	420
cgtctgtgcc acccagagcc ggccggccgc taggtccccgg gagaccctgc tatggtgctgt	480
gcgggcgcgg tgggggctca tctcccgcg tccggcttgg atatcttcgg ggacctgaag	540
aagatgaaca agcgccagct ctattaccag gttttaact tcgccatgtat cgtgtttct	600
gcactcatga tatggaaagg cttgatcgatgc acacaggca gtgagagccc catcggtgt	660
gtgctgatgt gcagtatgga gcccgcctt cacagaggag acctcctgtt cctcacaaat	720
ttccggaaag acccaatcag agctgagata atggagacat caaattctg actaaaggag	780
ataataatga agttgatgtat agaggcttgtt acaaagaagg ccagaactgg ctggaaaaga	840
aggacgttgtt gggaaagagca agagggtttt taccatatgt tggatggtc accataataa	900
tgaatgacta tccaaaattt aagtatgctc tttggctgt aatgggtgca tatgtgttac	960
taaaacgtga atcctaaaat gagaagcagt tcctggacc agattgaaat gaattctgtt	1020
gaaaaagaga aaaactaata tatttgagat gttccatgtt ctgtataaaa gggAACAGTG	1080
tggagatgtt ttgtcttgtt ccaaataaaa gattcaccag taaa	1124

<210> 33
<211> 2414
<212> DNA
<213> Homo sapien

<400> 33	
agactggctg aggaaggaat ttggggcaag agacaaaaat acagcaacag gagaaaagac	60

50

tcacggaggt agaaagagac tgggagacaa aaagagagaa acacatcaa aagatgtgga	120
gagagataga aacagagcca ggcagagtaa aaagaggctg agagagatga gtttagagatg	180
tgcagctgga catgttagagg acagagaaaa gcaaattggg ccagataatg tcaaagacct	240
tcaggcaaac ggagggcagc cagggagaca ggcgtgtgca cagcaaggct acagcctctc	300
ctgaccctgc cttccctcc ctactgtgga cgccaggagaa atccaaccca cacagtgaat	360
tcagccacca gaacctcatc atcaacacgc tctcgctttt ctttgctggc actgagacca	420
ccagcaccac tctccgctac ggcttccctgc tcatgctcaa ataccctcat gtcgcagaga	480
gagtctacaa ggagattgaa caggtggttg gcccacatcg ccctccagcg cttgatgacc	540
gagccaaaat gccatacaca gaggcagtca tccgtgagat tcagagattt gctgaccttc	600
tccccatggg tgtgccccac attgtcaccc aacacaccag cttctgaggg tacaccatcc	660
ccaaggacac ggaagtattt ctcatcctga gcactgctct ccgtgaccca cactacttg	720
aaaaaccaga cgccttcaat cctgaccact ttctggatgc caatgggca ctgaaaaaga	780
atgaagcttt tatccccttc tccttaggaa agcggatttg tcttggtaa ggcattgccc	840
gtgcggaatt gttccctttc ttcaccacca tcctccagaa cttctccgtg gccagccccg	900
tggctcctga agacatcgat ctgacaccccc aggagtgtgg tgtggcAAA ataccccaa	960
cataccagat ctgcttcctg ccccgctgaa ggggctgagg gaagggggtc aaaggattcc	1020
agggtcatTC agtgtccccca cctctgtaga taatggctct gactccctgc aacttcctgc	1080
ctctgagaga cctgctgcaa gccagcttcc ttcccttcca tggcaccagt tgtctgaggt	1140
cgcagtgcaa atgagtggag gagtgagatt attgaaaatt ataatataca aaattatata	1200
tatatatTTT gagacagagt ctcactcagt tgcccaggct ggagtgcagt ggcgtgatct	1260
cggctcactg caacctccac ccccggggtt caagaaaattc tcctgcctca gcctccctag	1320
tagctggat tacaggtgtg tgctaccatg cctggctaatttttattttt ttagtagaga	1380
tggggTTCA ccgtgttggc caggctgate tcaaactctt gaactcaagt gattcaccca	1440
ccttagcctc ccaaagtgc gggattacag gtgtgagtca ccatgcccg ccatgttatAT	1500
atataatTTT AAAAATTAAG atgaaattca cataaaataa aattagccat tttAAAGTGT	1560
acaattttagt ggtgtgtggt tcattcacaa agctgtacaa ccaccacat ctagttccaa	1620
acatTTTCTT TTTTCTGAG acggagtctc actctgtcac ccaggttcga gttcagtggT	1680
cttgaactcc tgatgtcagg tgattctctt agttccaaat gttttcatta tctctcccc	1740
aacaaaaccc atacctatca agctgtcact ccccataccc cattctctttt ttcatctcag	1800
ccccctgtcaa tctggTTTTT gtccttatgg acttaccaat tctgaatatt tcctataaaAC	1860
agaatcacac aatatttgat tttttttta aaactaagcc ttgctctgtc tcccaggctg	1920

gagtgtgtg gcgtgattt ggtaactgc aacctccgcc ttccaagttc aagagattct	1980
cctgcctcag ctccaagta gctggatta caggcatgtg gtaccacgcc tggctaattt	2040
tcttgatattt ttagtaggga catgtggcc aggctggttg tgagctcctg gcctcaggtg	2100
atccacacgc ctcagtgtcc cagagtgtg atattacagg cgtaatatgt gatctttgt	2160
gtctggttcc tttcacgttg aacgctattt ttgaggttcg tgccctgtgt agaccacagt	2220
cacacactgc tgttgtttc ccccatcctc attcccagct gcctcctcct actgtttccc	2280
tctatcaaaa agcctccttgcgcgaggttc cctgagctgt gggattctgc actggtgctt	2340
tggattccct gatatgttcc ttcaaattcca ctgagaattt aataaacatc gctaaagcct	2400
gacctccccca cgtc	2414

<210> 34
<211> 578
<212> DNA
<213> Homo sapien

<400> 34 atgctgctcg agcggcgcag tgtgtatggat ccggccgggc aggtacaaac ttatgaagaa	60
ggtctctttt atgctcaaaa aagtaagaag ccattaatgg ttattcatca cctggaggat	120
tgtcaataact ctcaaggact aaagaaagta tttgccccaaa atgaagaaat acaagaaatg	180
gctcagaata agttcatcat gctaaacctt atgcatgaaa ccactgataa gaatttatca	240
cctgatgggc aatatgtgcc tagaatcatg ttttagacc cttctttaac agtttagagct	300
gacatagctg gaagataactc taacagattt tacacatatg agcctcggtt tttaccccta	360
ttgatagaaa acatgaagaa agcattaaga cttattcagt cagagctata agagatgata	420
aaaaaaagcc ttcaattcaa agaagtcaaa tttcatgaag aaaacctctg gcacattgac	480
aaatactaaa tggcaagta tatagatttt gtaatattac tatttagttt tttatgtg	540
tttcaatag tcttattaaa ataaatgttt tttaatatc	578

<210> 35
<211> 1410
<212> DNA
<213> Homo sapien

<400> 35 tggctgtacg gcgagttta gatcctacgt ctggtccagt cggctttcct ccggccggg	60
ccctggccca gctagccggc catggaaggt aatggccccc ctgctgtcca ctaccagccg	120
gccagcccccc cgccccggacgc ctgcgtctac agcagctgct actgtgaaga aaatatttgg	180
aagctctgtg aatacatcaa aaaccatgac cagtatcctt tagaagaatg ttatgtgtc	240

52

<210> 36
<211> 734
<212> DNA
<213> *Homo sapien*

<400> 36
agagagagag agagagagag agagagagag agagtggaca taaaaattgc ttagtaaagg 60
tcaaaatgttc taaaactgcct gcatataagg atctcggtaa aaatctacca ttccctacat 120
attttcctgaa tggagatgaa gaggaactgc cagaagattt gtatgatgaa aacgtgtgtc 180
agccccgtgc gccttctatt acatttgctt aacatctttg gacgtggca gaaccttacat 240
attctgtgag ctgcgtatgag ccagagtgtat atcataaccat ccagaaatca tactctccctt 300
tcttagtcatcac aacaaaatca cacatgtcat ctttgtcaag ggcataaata tatcattcat 360
accccccatta aattttgtta gaaaaattac cacattaaat atatgagttt agtagattgg 420

53

atttgctgaa attgggtttg ggcataattttag caaaatattt ttaatttgtg gactcgattc	480
ttttttacta catatttccc aagttatctt aagatgtctg taaatttaac ttttattaaa	540
gttttgcata tctttgtgaa atagtgggtt tggaacagta gaaaaccata tggggactat	600
agtgcacacctt atttgggtaa agaaaccatt tgctaaaatg gagaaagtaa atagattttt	660
atttaaatta cagaaacatg ttaaaggccg gacaaaggaa agacaataaa atcataaattt	720
atcggtcctg ttta	734

<210> 37
<211> 683
<212> DNA
<213> Homo sapien

<400> 37 ggccatccag ccctgtggac cgaatggagt cccgcacgt gttgaggtca gttgtgggtt	60
ccccctggcct cggggctgggc gcggggtcag cgacactgca ggccgcgtt gcggtacggg	120
ctgggtgaaag tggagatgga cggcaggatg gattcacttg gccacatggc gcgaasstgg	180
gaagacggac accgaccta a gtcagtgta gtctaccact gtacatctgg taacctcaat	240
ccctgcaacc ggggcaaaat gggttccag gtcttggcaa cctttgaaat tccaattcca	300
ttttagagag ctttgacgag gcatatgct gattcacca ccagcaactt cagaacccag	360
tactggaatg ccatcagcca gcaggcccct gcatcatct atgacttcta tctgtggctc	420
actggaagga aacccaggca aggccaagat ggctcaaaga gcaaccagcc acctctgcag	480
cctgccacccctt cctgctggca agatttggtt ttgcattctg tgaagagcca aggaggcacc	540
agggcataag tctactact tatatctgta tggAACATAA cgcttggttt ttttacaac	600
aaataaaaatt gatcttgaat aaaaacagat gcggccggac atcctcatct atatttcgt	660
tcgacataaa tatgggtgta ttc	683

<210> 38
<211> 1181
<212> DNA
<213> Homo sapien

<400> 38 gcatgctgca acgactctct taatcctcca ccgtacaga ctaaatgagg gatttcttct	60
tggtttggat ccattgctgg caaagttgtt atctatgca caagccagag aaactgcagt	120
tcaacagtac aaaaaactgg aagaggaaat ccagaccctt cgagtttact acagtttaca	180
caaattttta tctcaagaag aaaatctgaa ggatcagttt aactataccc ttagtacata	240
tgaagaagct ttAAAAAAACA gagagaacat tgTTTCCATC actcaacaac aaaatgagga	300
actggctact caactgcaac aagctctgac agagcgagca aatatggaaat tacaacttca	360

acatgccaga	gaggcctccc	aagtggccaa	tgaaaaagtt	caaaagttgg	aaaggctggt	420
ggatgtactg	aggaagaagg	ttggaaccgg	gaccatgagg	acagtgtatct	gattaaaaaa	480
aaacgacagt	ctggggaaagc	gatcacatct	ggtgaccagg	ctgcttcatt	caacactgtg	540
taaacaccaa	agccttaact	tagcaaacag	ttgttagaag	tgggacactc	caaccacatt	600
ccaagctgag	ataaaatcaa	atcacaaatg	tttaaccact	ttgctgctga	cttgagttat	660
ttatccaaat	atattaacta	tagactttt	ccaatggta	gctataaggt	tacagcttat	720
tttgtaacta	ttttatatct	caatatcttt	aatataaatac	tttttactga	gagatcatta	780
tagaaacatg	ttaaagttgg	tttaggatcat	atcttcacat	atggcccttt	ctgaatcaaa	840
gtgcggcaaa	gtaaatattg	tctaagcttt	aatccactgt	gttaggtcaa	aacttcaaata	900
acatgcattt	ttcaatatacg	ggtatatttc	ttaactgatg	agagaggctt	agacatgagt	960
gtgtagtcctt	ccttcaatgc	gtgtatgtaa	tctttgttag	tataaaagat	attaaatata	1020
ggtgccaaga	attaaatgtt	taatttgttt	aataagagat	ggatataat	aaattacatt	1080
catcaaggca	tgatTTTGT	ttcactacaa	ataatgcaaa	ctgtttcaa	taaaaagagg	1140
agactgttaa	tgtgtactta	taaattcaca	ttgtcagtat	t		1181

<210> 39
<211> 2042
<212> DNA
<213> Homo sapien

<400> 39						
agtgatggcc	gccgtccccg	tgcaccccg	tgatggccgc	cgtccccgtg	cacaccagtg	60
atggccgccc	tgcccgtgca	ccccagtgtat	ggccgcgc	cccggtgcaca	ccagtgtatgg	120
cctctgtccc	ccatgcactc	ccagacaggc	aatgtccctg	tgggcctgtc	ccaggctctg	180
ttctcagcag	gctgggctca	gccctggtgc	agggagttag	gaggtgggag	tagtagggac	240
cagaaaaagt	ggcagctgtt	gacaactctg	ccatctcttt	ctgaatgtaa	tgggaggtcc	300
tgtctttca	gcttgcaagg	aaggagggtc	cgaggcaact	ccgctgttgc	acatttaggg	360
acccctgaac	ttaaatgaca	gaatgcctg	accactctgg	aaggcactgt	gttcatgttt	420
gtgtgcttga	ctcttgatcc	gtaaaatggc	tgtttgtgca	ggtcattaac	tgtgagattc	480
agagagtagg	tgcacacgtc	cctgcagaga	ttccagcagg	actgaaaacc	agtagaaata	540
tatcagcacc	tggatcttgc	ctcctgagtc	agtaaggata	tgccacagtc	acgaaggcag	600
tgggatttgc	agggagggaa	gggaaggcgg	caggcggggc	atgcctccg	gggtgcccga	660
acacacctgc	tgcattccaca	tgtcttcaga	gccctctccc	tgtggaggc	cttttcagg	720
acagccttgg	tgaactggaa	acggaatccc	agcccttggt	ggccctgcag	tgacttggac	780

ctttccgagg tcaccctgcc	actgcgtgcc	cttcagtc	cc tcctggcagg	tgggggcaca	840
tcccccagcc actcccattt	cctgacattt	tcactttgt	taactggaag	ccttctgtga	900
aattttagtt ttc	aaagcat	tatctggtga	tggcaaccc	agggcagcga	960
attttcttat	ctaggctaat	aaacataata	aaatcaataa	ggactttgaa	1020
ctgggttcag	gaaactgagt	gtggccgccc	tgtgggtgg	tgttttgtga	1080
gaggtgagta	gttaattcac	aggagtgact	aatggcagcg	tcccactcac	1140
ggggtcatgg	tctcaagggg	tcactccatg	cactggggat	gtcagctcat	1200
tatattcggg	aagtgtctca	gttctgagtg	ccttgaggg	aatttgact	1260
cacagccttg	cattgtgtgt	gttagaggct	gtgggccttg	ggcaggaggg	1320
gcacatacct	cccgctctc	ccagcctct	ctgactctga	ctttccctct	1380
cggctctctg	accagttcca	cgacatcctc	attcgaaagt	ttgacaggca	1440
cagatcgct	tcgacgactt	catccagggc	tgcacatgtcc	tgcagaggtt	1500
ttcagacggt	acgacacgga	tcaggacggc	tggattcagg	tgtcgtacga	1560
tccatggtct	tca	gtatcgatgt	atgaccctgg	cctctcgta	1620
gagccaaaat	gtcacagttc	ctatctgtga	ggaaatggag	cacaggtgca	1680
gttctccctt	tagatttgt	cacgtggga	cccagctgt	catatgtgga	1740
aatggttttg	caactgtaat	agtagctgt	tcgttcta	atgcacattt	1800
ctgtctcatt	gtgccatgag	gtaaatgtaa	tgtttcaggc	attctgcttgc	1860
tatcatgtgc	ttttcttagat	gtctctgggtt	ctatagtgc	aatgc	1920
ggaattttaa	aataacatgg	aacttacaca	aaaggctttt	catgtgcctt	1980
aaaggagttt	attgttattca	ttgaaatatg	tgacgtaagc	aataaaggga	2040
tg					2042

<210> 40
<211> 1287
<212> DNA
<213> Homo sapien

<400> 40	ggtgataatg	ccaggccctg	ccccggcag	aggcggaa	ggagtcggcc	tgagagg	tct	60
	ctcgtegctg	caggcgcc	agcccagccg	cgtgc	ccttgg	cccatggcc	cctactctt	120
	ccgccccggc	cctggggccg	gccttggcc	tgctgc	aggc	gcggcg	ctgc	180
	cttcctgtgg	aacgtttcc	agagggtcga	taa	agacagg	agtgg	agtgtga	240
	cgagcttcag	caagctctc	ccaacggc	gtggactccc	tttaatccag	tgactgtc	ag	300

gtcgatcata tccatgtttg accgtgagaa caaggccggc gtgaacttca gcgagttcac	360
gggtgtgtgg aagtacatca cggactggca gaacgtttc cgcacgtacg accgggacaa	420
ctccggatg atcgataaga acgagctgaa gcaggccctc tcaggtttcg gctaccggct	480
ctctgaccag ttccacgaca tcctcattcg aaagtttgcac aggcaggac gggggcagat	540
cgccttcgac gacttcatcc agggctgcat cgtcctgcag acccttgctc catcacccag	600
gccagagtgt ggtggcgcga acacggctca ctgcagcctc gaccctcagg ctcaagcgat	660
cctcacgcct cgAACCCCCCA aagtgttggg atcacaggcg agagtcacca tgctggcctg	720
aatcttcagg aggttgacgg atatattcag acgttacgac acggatcagg acggctggat	780
tcaggtgtcg tacgaacagt acctgtccat ggtcttcagt atcgatgac cctggcctct	840
cgtgaagagc agcacaacat ggAAAGAGCC aaaatgtcac agttcctatc tgtgaggaa	900
tggagcacag gtgcagttag atgctgttct tccttttagat tttgtcacgt ggggacccag	960
ctgtacatat gtggataagc tgattaatgg ttttgcact gtaatagtag ctgtatcgat	1020
ctaattgcaga cattggattt ggtgactgtc tcattgtgcc atgaggtaaa tgtaatgttt	1080
caggcattct gcttgcaaaa aaatctatca tgtgcTTTC tagatgtctc tggttctata	1140
gtgcAAATGC ttttatttagc caataggaat ttAAAATAA catggaaacctt acacaaaagg	1200
cttttcatgt gccttacttt ttAAAAAAGG agtttattgt attcattgga atatgtgacg	1260
taagcaataa agggaaatgtt agacgtg	1287

```
<210> 41  
<211> 1763  
<212> DNA  
<213> Homo sapien
```

<400> 41
aaaaagatca gagcgcagcc gaggacccgg cgagagcaag gacgcgcgct cggcgacgca 60
gcgcaagga acacaataca caccgagcat gtaaggccgc cgcgcgcc ccacacgcgt 120
acccagcaca tacgggtgcag agaggacgac gtggccgtcc acaccccgta gcaccagcca 180
acgccccgca cctcggccctc tctctgattt ccttatgtgt tgttgttact ttgtttgtta 240
ttgtttgttc tgtgattgtt tgttattttt atttattatt ttgttttgtt gttgtttgtg 300
tttttgtgtt ttgtttttt ttgtttttt tgttttttt tttttttaat ttttgtattc 360
ttataaatgt gtttaattac aactgcttca aaagaatccc agctttcaa aagtttattt 420
taagtttggaa gactagacaa ggtcatactg gtttacatc ctacgtgata taagtatata 480
tacaaaagaaa aaaacaacat tggaatatta cacagcttga aggtttgcaa aggttatttg 540
tgtcttagtt atttctgcac ttaatgacac atcagacgca ttgagttatata ttcataagtt 600

gttgactagc aaagatacaa tcattagtaa cccaaagtctt caaaaattcac accaaacttt	660
atgaagtcat tcagaaagag aaagtcaatc ctaaaattaa aattggcaac tatgataaat	720
accttcaaaa ggatgttagat ataatggaga tggtaaaag tttagttca ttaattgtaa	780
aattagcatg ttatatttac tcaatatagt gaagactagg tgattcttac atgtattcta	840
cttatggtagt gttactggtt ttagtgtgaa ttacataga ataaatttac ttcactttca	900
tgtcatcgac atgaatgaca caaaagctac ttcataatac tactttacaa tagtttcaa	960
catttccata tggtgcgacc ctttgcctt catcaatttt ggggtgtcatg agaacaatag	1020
gtatcccgtt ggacatgtatg tattgcgaag agcatataaa gcagagggaa aatgaaaaag	1080
caagagaaac tcatttcaat gcttttcta aaaggtaaca aatataattt taatcaactt	1140
ccttgaaaaa tatttttaaa acaggtatca atagaaaaaa ttacaaaaca tcataatgaag	1200
ctataaataa tttgaaaaaa ctatatcatc ataaagcata agtaataatc ttaaaaatac	1260
actcttaaga aggtatgtaa ttgc当地 aaaaatggcta gatatctgtat gggacagtaa	1320
accttggaaag aaactggcta aagagtaagt gtgtgtatat ttctgaacct aagtaattat	1380
ttgtcacgac tttaaaattt agccagttac aaatattttt aaatcctaac tttaaagtta	1440
tctaaaaaaag gcaatatgga ggaaatagta attttgc当地 tgaaaatgt tgaaaactga	1500
tcaccatttc agaggcttca aattcataat ttcatatataa gaacaagaag tagaaagcat	1560
atggggcaagg aacaaatatg tggccagcca gtc当地 acgaactaat ttgttctta	1620
ttaaaaatgc caatacataatt gactttctt ttaaatttctt cactatgatt gaagaccact	1680
ccatatatac atcattaaga aatgctgtta acacatggac agacaagaca gtaacagtct	1740
agtggctttt gttatgcagc aca	1763

<210> 42
<211> 2913
<212> DNA
<213> Homo sapien

<400> 42	
cccgtaggg gttaccctt ccatttaag caggatattc tagatctct cagtctcaca	60
gccttgc当地 ccaataacca gcagaaagcg gttcgacaat tggcccttct ttggccct	120
cctgc当地 cccggattt gacggctgag tctggctacg cggccctccg cgggagcg	180
atggggccaa tcaagagctt ggc当地 tttt acaaactgag aaagtagctc cagcagcacc	240
cgagagggtc aggagaaaag cggaggaagc tggtaggccc ctgaggggccc tcggtaagcc	300
atcatgacca cccggcaagc cacgaaggat cccctctcc ggggtgtatc tcctaccct	360
agcaagattt cggtagctc tcagaaacgc acgc当地 ccactgttac atcgtgc当地	420

gtggaccagg agaaccaaga tccaaggaga tgggtgcaga aaccaccgct caatattcaa	480
cgcgcctcg ttgattcagc aggccccagg ccgaaagcca ggcaccaggc agagacatca	540
caaagattgg tggggatcag tcagcctcg aacccttgg aagagctcag gcctagccct	600
aggggtcaaa atgtggggcc tgggccccct gcccagacag aggctccagg gaccatagag	660
tttgtggctg accctgcagc cctggccacc atcctgtcag gtgagggtgt gaagagctgt	720
cacctggggc gccagcctag tctggctaaa agagtactgg ttcgaggaag tcagggaggc	780
accacccaga gggtccaggg tgttcgggcc tctgcatatt tggcccccag aaccccccacc	840
cacccactgg accctgccag ggcttcctgc ttctcttaggc tggagggacc aggacctcga	900
ggccggacat tgtgccccca gaggctacag gctctgattt cacctcagg accttccttt	960
cacccttcca ctcgccccag tttccaggag ctaagaaggg agacagctgg cagcagccgg	1020
acttcagtga gccaggcctc aggattgctc ctggagaccc cagtcagcc tgctttctct	1080
cttcctaaag gagaacgcga ggttgtcaact cactcagatg aaggaggtgt ggcctctctt	1140
ggtctggccc agcgagtacc attaagagaa aaccgagaaa tgtcacatac cagggacagc	1200
catgactccc acctgatgcc ctccccctgcc cctgtggccc agcccttgcc tggccatgtg	1260
gtgccatgtc catcacccctt tggacgggct cagcgtgtac cctccccagg ccctccaact	1320
ctgacacctat attcagtgtt gcggcgtctc accgttcaac ctaaaacccg gttcacaccc	1380
atgccatcaa cccccagagt tcagcaggcc cagtggtcgc gtgggtctc ccctcagtcc	1440
tgcctctgaag atcctgcccct gcccctggag caggttgcgc tccgggtgtt tgaccaggag	1500
agttgtataa ggtcactgga gggttctggg aaaccacccg tggccactcc ttctggaccc	1560
cactctaaca gaaccccccag cctccaggag gtgaagattc aagtgagtct gtgtggccaa	1620
cagctttgat gtctattgaa cagtgactgg gctgaggaag agggaaaaga gatggggat	1680
caggaatagg acagtgtggg tagactactg aacgcacatc ttgatgtcac actggggtgc	1740
tctctccac cacagcgcac cggtatcctg caacagctgt tgagacagga agtagagggg	1800
ctggtagggg gccagtgtgt ccctcttaat ggaggcttt ctctggatat ggttgaactt	1860
cagccctgc tgactgagat ttctagaact ctgaatgcca cagagcataa ctctggact	1920
tcccacccctc ctggactgtt aaaacactca gggctgccaa agccctgtot tccagaggag	1980
tgcgaaaaac cacagccctg ccctccggca gagcctggc ccccgaggc cttctgttagg	2040
agtgagccctg agataccaga gcccctccctc caggaacacgc ttgaagtacc agagccctac	2100
cctccagcag aaccaggcc cctagagtcg tgcgttagga gtgagcctga gataccggag	2160
tcctctcgcc aggaacagct tgaggtaccc gagccctgcc ctccagcaga acccaggccc	2220

59

ctagagtccct	actgttaggat	tgagcctgag	ataccggagt	cctctcgcca	ggaacagctt	2280
gaggtaacctg	agccctgcc	tccagcagaa	cccgggcccc	ttcagccag	cacccagggg	2340
cagtctggac	ccccagggcc	ctgccttagg	gtagagctgg	ggcatcaga	gcctgcacc	2400
ctggAACATA	gaagtctaga	gtccagtcta	ccaccctgt	gcagtca	ggctccagca	2460
accaccagcc	tgatcttc	ttccaaacac	ccgctttgt	ccagcccccc	tatctgctca	2520
ctccagtctt	tgagacccccc	agcaggccag	gcaggcctca	gcaatctggc	ccctcgaaacc	2580
ctagccctga	gggagcgcct	caaatcgtgt	ttaaccgcca	tccactgctt	ccacgaggt	2640
cgtctggacg	atgagtgtgc	cttttacacc	agccgagccc	ctccctcagg	ccccacccgg	2700
gtctgcacca	accctgtggc	tacattactc	aatggcagg	atgcctgtg	tttatttcca	2760
gttggttctg	ctgccccccca	gggctctcca	tgatgagaca	accactcctg	ccctgcccgt	2820
cttcttc	ttagccctta	tttattgtcg	gtctgccc	ggactggga	gccgcccact	2880
tttgtcctca	ataaaagttt	taaaagtaaaa	cac			2913

<210> 43
<211> 986
<212> DNA
<213> Homo sapien

<400> 43						
cgccaggaac	agcttgaggt	acctgagccc	tgccagctcc	agcagcaccc	gagagggtca	60
ggagaaaaagc	ggaggaagct	gggttagggcc	tgaggggccc	cggttaagcca	tcatgaccac	120
ccggcaagcc	acgaaggatc	ccctcctccg	gggtgtatct	cctacccct	aggtagagc	180
tggggccatc	agagecctgc	accctgaaac	atagaagtct	agagtccagt	ctaccaccc	240
gctgcagtca	gtgggctcca	gcaaccacca	gcctgatctt	cttttcccaa	cacccgctt	300
gtgccagccc	ccatatctgc	tcactccagt	cttgagacc	ccagcaggc	caggcaggta	360
aggagttggc	tgggaaggag	tgtgaacaca	agaggtcctc	acctcactgt	gagctgcaca	420
cctgccc	ccctacccca	ggcaatctca	tgcttccaca	ccttccaccc	tggccagcc	480
tggctctccc	tcaggaagag	gggaggggct	gcacttccag	ccctgtgctc	ctaattggct	540
tggccgttgg	tggggagga	ggagaggaca	gtacatggtg	gaagtatagg	accccagacc	600
tccctctaaa	ttttccatgc	ccctcaggcc	tcaaatct	ggccctcgt	accctagccc	660
tgagggagcg	cctcaaattcg	tgtttaacccg	ccatccactg	cttccacgag	gtctgtctgg	720
acgatgagtg	tgcctttac	accagccgag	cccctccctc	aggccccacc	cgggtctgca	780
ccaaaccctgt	ggctacatta	ctcgaatggc	aggatgcct	gtgtttcatt	ccagttggtt	840
ctgctgcccc	ccagggctct	ccatgatgag	acaaccactc	ctgcctgccc	gtacttcttc	900

60

cttttagccc ttatTTATTG tcggTctGCC catGGGACTG ggAGCCGCC	actTTTGTCC	960
tcaataaaagt ttctaaAGTA aaACAC		986

<210> 44
<211> 865
<212> DNA
<213> Homo sapien

<400> 44		
ccctgctgat acgattcgag ctcgtacccc tccagctggc cccaaggaga aagccttctc		60
aagttagata gaagatttgc cgtacccccc caccacagaa atgtatttgt gtcgttggca		120
ccagcctccc ccatcaccgt taccattacg ggaatcctct ccaaagaagg aggagactgt		180
agcaagtaag gcatagagaa cacttgctct tataccctag tggtggcggt caagctaaca		240
agtgtgaaaa tgccttggc atttttaaaa aagtgcatac aataaagcag agttctgtca		300
agaatggata agttaacagc cagagacaga cactgtgcag gcattgcaaa tagatggat		360
tacagcaaaa tgtgctaat gtatTTGCCT gtttacaaca ctgggagatg tgTTTGCAG		420
taagttgctc atcacaagag caccagactt ggggtgtaa tctccggcaa cttgcattgcc		480
ctctgaaaga agggTTTCT gtgtgtgaa atgcatacgaa ctatactttg ccatgcacga		540
ctgttccTgc aattgatatt gtgtgaaatc tgggaggggtg gtctttgggt gtttcaggg		600
gccaatggta atttttgggt tggggagcca gcttgggggtg gggaaTTTC acctgggcct		660
ccgctttta actatataaa catttatctg tatatctatg tccctgtctg gggggcagga		720
ggaatctgcc aaagaccaac agtcttactt tattttacta tacttcacaa aggttctaaa		780
atgtgaagag ttacttggg ttgcagtagc ccattgggtt ttcatatatt taaataaaat		840
ggtctacaaa ctatTTTCA aacaa		865

<210> 45
<211> 1050
<212> DNA
<213> Homo sapien

<400> 45		
cccccgccgc cctcgctccc tcccgtcagc ccccgccccct cggcgaagg agcggcgtgc		60
cgtccgggtc gcctaggcct ggggtcgggc gcgccacgc tgcgtgcctt gggcgcgctc		120
gggattctcg cctggcgcgg ctggggagg tgaacagtgt ggcccgcatt gttttctcc		180
gcggcgctcc gggccccggc ggctggcctc accgcccact ggggaagaca tgtaaggaat		240
ttgcataaga cagctatgca aaatggagct ggaggagtt tattttgtca cagagatact		300
cctgagaata accctgatac tccatttgat ttcacaccag aaaactataa gaggatagag		360
gcaattgtaa aaaactatcc agaaggccat aaagcagcag ctgttcttcc agtcctggat		420

tttagccaaa ggcagaatgg gtgggtgccc atctctgcta tgaacaaggt tgcagaagtt	480
ttacaagtac ctccaatgag agtataatgaa gtagcaactt tttataacaat gtataatcga	540
aagccagttg gaaaagtatca cattcaggc tgcaactacta caccctgcat gcttcgaaac	600
tctgacagca tactggaggc cattcagaaa aagcttgaa taaaggttgg ggagactaca	660
cctgacaaac ttttcaactt tatagaagtg gaatgtttag gggcctgtgt gaacgcacca	720
atggttcaaa taaatgacaa ttactatgag gatttgacag ctaaggatat tgaagaaatt	780
attgatgagc tcaaggctgg caaaatccca aaaccaggc caaggagtgg acgcttctct	840
tgtgagccag ctggaggctt tacctctttg actgaaccac ccaagggacc tggatttgg	900
gtacaatgtg ttcaccccca cagggaaattc caaggtgcaa tagcggttgt tgtcaatcat	960
aggatctctg ttgggatggc tgaaggtgaa acagggctgg ggtgcggaga gctgggtggaa	1020
gttgcgcagc cgtacctgcc cggggcgccg	1050

<210> 46
<211> 1027
<212> DNA
<213> Homo sapien

<400> 46	
ccccgcgcgc cctcgctccc tcccgctcagc ccccgccctt cggcgaagg agcggcgtgc	60
cgtccgggtc gcctaggcct ggggtcgaaa ggcgcacgc tgtgcgcctt gggcgcgtc	120
gggattctcg cctggcgccgg ctggggaaagg tgaacagtgt ggcccgccat gttcttctcc	180
gcggcgctcc gggccgggc ggctggcctc accgcccact gggaaagaca tgtaaggaat	240
ttgcataaga cagctatgca aaatggagct ggaggagctt tatttgcata cagagatact	300
cctgagaata accctgatac tccatttgat ttcacaccag aaaactataa gaggatagag	360
gcaattgtaa aaaactatcc agaaggccat aaagcagcag ctgttcttcc agtcctggat	420
tttagccaaa ggcagaatgg gtgggtgccc atctctgcta tgaacaaggt tgcagaagtt	480
ttacaagtac ctccaatgag agtataatgaa gtagcaactt tttataacaat gtataatcga	540
aagccagttg gaaaagtatca cattcaggc tgcaactacta caccctgcat gcttcgaaac	600
tctgacagca tactggaggc cattcagaaa aagcttgaa taaaggttgg ggagactaca	660
cctgacaaac ttttcaactt tatagaagtg gaatgtttag gggcctgtgt gaacgcacca	720
atggttcaaa taaatgacaa ttactatgag gatttgacag ctaaggatat tgaagaaatt	780
attgatgagc tcaaggctgg caaaatccca aaaccaggc caaggagtgg acgcttctct	840
tgtgagccag ctggaggctt tacctctttg actgaacggc ctccagttatg ctgtcagat	900
ttcgaagcat gcagggtgtt gtagtgcaga cctgaatgtg atactttcca actggcttcc	960

gattatacat	tgtataaaaa	gttgctactt	cataactct	cattggaggt	acctgcccgg	1020
gcggcccg						1027

<210> 47
<211> 864
<212> DNA
<213> Homo sapien

<400> 47						
ccccgcgcgc	cctcgctccc	tcccgtcagc	ccccggccct	cggcgaagggg	agcggcgtgc	60
cgtccgggtc	gcctaggcct	ggggtcggga	gcgcgcacgc	tgtgcgcctt	gggcgcgctc	120
gggattctcg	cctggcgccgg	ctggggaaagg	tgaacagtgt	ggcccgcoat	gttcttctcc	180
gcggcgctcc	ggggccgggc	ggctggccctc	accgcccact	gggaaagaca	tgtaaggaat	240
ttgcataaga	cagctatgca	aatggagct	ggaggagctt	tatttgtca	caagagatact	300
cctgagaata	accctgatac	tccatttgat	ttcacaccag	aaaactataa	gaggatagag	360
gcaattgtaa	aaaactatcc	agaaggccat	aaagcagcag	ctgttcttcc	agtcctggat	420
ttagccaaa	ggcagaatgg	gtggttgccc	atctctgcta	tgaacaaggt	tgcagaagtt	480
ttacaagtac	ctccaatgag	agtatatgaa	gtagcaactt	tttatacaat	gtataatcga	540
aagccagttg	gaaagtatca	cattcaggtc	tgcactacta	caccctgcat	gttgc当地	600
tctgacagca	tactggaggc	cattcagaaa	aagcttggta	ggaaatacat	gatatttgt	660
acactgataa	aaagtataat	tgtctctcta	gatttggtag	atttctatct	aaaatttcca	720
acttctgcca	tcttatttgg	tctgtactta	cctagtaata	ttttgtgtta	ctgtgtttcc	780
acatctttat	ttcttcctat	tttgtattct	tcctcagttc	ttagtgttaa	agctgagttt	840
ttaatttttt	cttttttaat	cagt				864

<210> 48
<211> 1014
<212> DNA
<213> Homo sapien

<400> 48						
gagcggcgca	gtgtgatgga	ttcgccggcg	aggtacatcc	ccttcaagca	gtatgtggc	60
aaatacgtcc	tcttgcac	gtggccagct	actgaggcct	gacggccag	tacatcaagt	120
ggatcgtctc	tgccccctt	gcccaggta	gcgcaggttt	ctttgtccctg	gggagccggg	180
cgcgaagagc	ggcgctatc	tctcggagg	tgtacccct	tatactgagt	gtgaccacgc	240
tcagcctctt	gctcgccccg	gtgctgtgga	gagctgcaat	cacgagggt	gtgcccagac	300
cggagagacg	gtccagcctc	tgtatggctcg	gagatgatgg	accgtggaag	ggaagcgtct	360

63

gtggggagtg agcgcttaga tggccacgag ctgctccttc tgggaagctc gcaccccttgc	420
aacagaacag ccctctagca gagcgtcagt gcagtcgtgt tatcccgct tttacagaat	480
attcttgc tatttttagaa tttccggag tagtttattt gcagtctgtt gattatgtgc	540
agtagacccg ggacactgcg ttttaccat caccttgaat gtgggcctg gatgtgcctt	600
ttttttttt ccctgaaatt attattaatt ttctattgtg agttcatcag ttcatagttt	660
tttttagtaaa gaagcaaaaat taaaaggctt taaaaatgt acaacttcag aattataatc	720
tgttagtcaa atatttgtta ttaaacattt ctgtaatatg aagttgtaat cctggccgtg	780
agcttggaaag cttacttttgc attcttaaag cctatgtttt ctaaaatgag acaaatacgg	840
atgtctattt gccttttattt gtaacttttta aatgaaataa tttcatgtca atttctatta	900
gatatatcac ttaaaatattt tggttttaaa tcacaagaat atgtatttctt taataaagat	960
aatttatgtat catggtataaa ttaattgaaa tttatttaaaa tctgttttta ttaa	
	1014

<210> 49
<211> 1509
<212> DNA
<213> Homo sapien

<400> 49	
ggtccaacgc cagccctgcgg ctgccaggcc ccacgcccgc caggaagtgc tgcgcgcgg	60
cggccgacgg gacccggccca cgccccgcct cttaaagggg gcagtgactg cggctggcg	120
ggagtccggg tcggcttggc tgagcggggg cgggtctggg cagggcggcg gccgcctcc	180
cccggaactcc cggccctcccg gcctccctgg tcccgctgg gaaggatgc aaggaagccc	240
tccggcgctg cgctccgagg cgggagacag cgtccccctc cggccctcgg gtcctggcg	300
ctcagagccc ggcccaggcc gcggAACGGT gatgctcggg cggacgggc gggcgccggat	360
ccctgcgtcc cgctgaaaat gtgtgtctga catgcaagct cagtgggca gagaccgtg	420
gattgctgtg ccctgcctc cggacctgga tcatgaaggt gttggaaaga agcttcttct	480
gggtgctgtt tcccgccctt ccctggcg tgcaggctgt ggagcacgag gaggtggcg	540
agcgtgtat caaaactgcac cggggcgag ggggtggctgc catgcagagc cggcagtgg	600
tccgggacag ctgcaggaag ctctcaggcc ttctccgcca gaagaatgca gttctgaaca	660
aactgaaaac tgcaattgga gcagtggaga aagacgtggg cctgtcgat gaagagaaac	720
tgtttcaggt gcacacgttt gaaatttcc agaaagagct gaatgaaagt gaaaattccg	780
ttttccaagc tgtctacgga ctgcagagag ccctgcaggg ggattacaat gatggaccgt	840
ggaagggaag cgtctgtggg gagtgagcgc ttagatggcc agcagctgct cttctggga	900
agctcgacc ttggcaacag aacagccctc tagcagagcg tcagtcgtcgt cgtgttatcc	960

64

cggttttac agaatattct tgccttattt tagaattttc cggagtagtt tatttcgagt	1020
ctgttgatta tgtgcagtag accccggaca ctgcgtttta ccgatcacct tgaatgtgg	1080
gcctggatgt gcctttttt ttttccctg aaattattat taattttcta ttgtgagttc	1140
atcagttcat agttttttta gtaaaaaggc aaaattaaaa ggctttaaa aatgtacaac	1200
ttcagaatta taatctgtta gtcaaataatt tgtttattaa catttctgtt atatgaagtt	1260
gtaatcctgg ccgtgagctt ggaagcttac ttttgcattt taaagcctat gttttctaaa	1320
atgagacaaa tacggatgtc tatttgcctt ttattgttac tttaaatgtt aataattca	1380
tgtcaatttc tatttagatat atcacttaaa atatttggtt tttaaatcaca agaataatgtt	1440
tttttaata aagataattt atgatcatgg tataattaat tgaaattttt taaaatctgt	1500
ttttattaa	1509

<210> 50
<211> 1206
<212> DNA
<213> Homo sapien

<400> 50	
ggtccaacgc cagcctgcgg ctgccaggcc ccacgcccgc caggaagtgc tcgccccccg	60
cggccgacgg gacccgccc cgcggccct cttaaagggg gcagtgactg cggctggcg	120
ggagtcgggg tcggcttggc tgagcgggggg cggtgctggg cagggcggcg gcccgtccct	180
cccgactcc cggcctcccg gcctccctgg tcccgctgg gaaggatgc aaggaagccc	240
tccggcgctg cgctccgagg cggagacag cgtccccctc cggccctcgg gtcctggcg	300
ctcagagccc ggcccaggcc gggaaacggt gatgctcggg cggacgggc gggcgccgat	360
ccctgcgtcc cgctgaaaat gtgtgtctga catgcaagct cagtggggca gagaccctgt	420
gattgtgtg ccctgcctc cggacctgga tcatgaaggt gttggaaaga agcttcttct	480
gggtgtgtt tcccgctt ccctggcg tgaggctgt ggagcacgag gaggtggcg	540
agcgtgtat caaactgcac cggggcgag ggggtggctgc catgcagac cggcagtgg	600
tccggacag ctgcaggaag ctctcagggc ttctccgcca gaagaatgca gttctgaaca	660
aactgaaaac tgcatttggc gcagtggaga aagacgtggg cctgtcggat gaagagaaac	720
tgttcaggt gcacacgttt gaaattttcc agaaagagct gaatgaaagt gaaaattccg	780
ttttccaagc tgtctacgga ctgcagagag ccctgcagggg ggattacaaa gatgtcgtga	840
acatgaagga gagcagccgg cagcgcctgg agggccctgag agaggctgca ataaaggaag	900
aaacagaata tatggaaactt ctggcagcag aaaaacatca agttgaagcc cttaaaaata	960
tgcaacatca aaaccaaagt ttatccatgc ttgacgagat tcttgaagat gtaagaaagg	1020

65

cagcggatcg tctggaggaa gagatagagg aacatgcttt tgacgacaat aaatcagtaa	1080
gcgttccaga acagctgctt cttcacctcc tgagccactc actaatcaga agacatgtt	1140
ttgaaattgt tcacgtgtat gtttttaatg tagattgaaa atgaagacaa actaaaatgc	1200
ttctct	1206

<210> 51
<211> 882
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (43)..(43)
<223> n=a, c, g or t

<400> 51	
tggtaattg gattctcacc cctccgcct acgcactgca ctncgactct tagagatccc	60
cggggagccg gggcagacgt ccgtagegcc ccctcccgag gaggtcgagc cgggcagtgg	120
ggtccgcata cttgtggagt actggtgagc ggccccggct ggaggacccg caccctggtc	180
ccgcgggccc gacggagggtg ggtccacggg aggccccacc cccgaatccc cagcccagcc	240
ccatcttttgc actccccagt gaaccctgcg gcttcgaggc gacctacctg gagctggcca	300
gtgctgtgaa ggagcagtat ccgggcatacg agatcgagtc ggcgcctcggg ggcacaggtg	360
cctttgagat agagataaat ggacagctgg tttctccaa gctggagaat gggggcttc	420
cctatgagaa agatctcatt gaggccatcc gaagagccag taatggagaa accctagaaa	480
agatcaccaa cagccgtcct ccctgcgtca tcctgtact gcacaggact ctgggttcct	540
gctctgttct ggggtccaaa ctttgtctc ctttgtcc tgctggagc tccccctgcc	600
tctttccct acttagctcc ttagcaaaga gaccctggcc tccactttgc ctttggta	660
caaagaagga atagaagatt ccgtggcctt gggggcagga gagagacact ctccatgaac	720
acttctccag ccacctcata cccccctccc aggtaagtg cccacgaaag cccagtcac	780
ttttcgcctc ggtataatccct gtctgatgcc acagatttttta ttattctcc cctaaccag	840
ggcaatgtca gctattggca gtaaagtggc gctacaaaca ct	882

<210> 52
<211> 1074
<212> DNA
<213> Homo sapien

<400> 52	
taaatgaagc catgaagtcc agcggacacc gggagtgggg agtggggaaag cccggcactc	60
cgggagacccg ggcaggaa ggaggctg gaccggaccc agccctgcc cggggagcga	120

gctccggagc	tgcctacga	ggtaaaaacg	tagcagtggc	ggagacccgc	aggggcgcc	180
cgaacgccac	cctcgcccc	tccccgtcc	agaggccccg	ccccgtcacg	tgcccgcggt	240
tcgcgtcaca	cccggaagca	ggggcccgag	cggaccggcc	gcgatgagcg	gggagccggg	300
gcagacgtcc	gtagcgcccc	ctcccgagga	gtcgagccg	ggcagtgggg	tccgcacatgt	360
ggtggagtac	tgtgaaccct	gcggcttcga	ggcgcacctac	ctggagctgg	ccagtgtgt	420
gaaggagcag	tatccgggca	tcgagatcga	gtcgccctc	gggggcacag	gtgccttga	480
gatagagata	aatggacagc	tggtgttctc	caagctggag	aatgggggt	ttccctatga	540
gaaagatgtg	agtatttaca	gcgttggag	gacctttgg	tcaccctacc	ccaacagtgc	600
atcatcctgt	cattccactc	ctctagctca	ttgaggccat	ccgaagagcc	agtaatggag	660
aaaccctaga	aaagatcacc	aacagccgtc	ctccctgcgt	catcctgtga	ctgcacagga	720
ctctgggttc	ctgctctgtt	ctggggtcca	aaccttggtc	tccctttgg	cctgctggga	780
gctccccctg	cctctttccc	ctacttagct	ccttagcaa	gagaccctgg	cctccacttt	840
gccccttggg	tacaaagaag	gaatagaaga	ttccgtggcc	ttgggggcag	gagagagaca	900
ctctccatga	acacttctcc	agccacactca	taccccttc	ccagggttaag	tgcccacgaa	960
agcccagtcc	actcttcgcc	tcggtaatac	ctgtctgatg	ccacagattt	tatattattct	1020
cccctaacc	agggcaatgt	cagctattgg	cagtaaagtg	gchgctacaaa	cact	1074

<210> 53
 <211> 961
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (43)..(43)
 <223> n=a, c, g or t

tgggtattg	gattctcacc	cctccgcct	acgcactgca	ctncgactct	tagagatccc	60
cggggagccg	gggcagacgt	ccgtagcgcc	ccctcccgag	gaggtcgagc	cgggcagtgg	120
ggtccgcata	gtggtgagt	actggtgagc	ggcccccgt	ggaggacccg	caccctggtc	180
ccgcggccg	gacggaggtg	ggtccacggg	aggccccacc	cccgaatccc	cagcccaagcc	240
ccatctcttg	actccccagt	gaaccctgcg	gcttcgaggc	gacctacctg	gagctggcca	300
gtgctgtgaa	ggagcagtat	ccgggcatacg	agatcgagtc	gcgcctcgaa	ggcacaggtg	360
cctttgagat	agagataaat	ggacagctgg	tgttctccaa	gctggagaat	gggggcttcc	420
cctatgagaa	agatgtgagt	atttacagcg	ttgggaggac	ctcttggtca	ccctacccca	480

acagtgcatc atcctgtcat tccactcctc tagtcattg aggccatccg aagagccagt 540
 aatggagaaa ccctagaaaa gatcaccaac agccgtcctc cctgcgtcat cctgtgactg 600
 cacaggactc tgggttcctg ctctgttctg gggtccaaac cttggtctcc ctttggtct 660
 gctgggagct cccccctgcct ctttccccta cttagtcct tagcaaagag accctggcct 720
 ccactttgcc ctttgggtac aaagaaggaa tagaaagattc cgtggccttg gggcaggag 780
 agagacactc tccatgaaca cttctccagc cacctcatac ccccttccca gggtaagtgc 840
 ccacgaaagc ccagtccact cttcgctctg gtaatacctg tctgatgcc aagatttat 900
 ttattctccc ctaaccagg gcaatgtcag ctattggcag taaagtggcg ctacaaacac 960
 t 961

<210> 54
 <211> 1839
 <212> DNA
 <213> Homo sapien

<400> 54
 ggagagatcg tccaggaggc ggtgttgatg cggcaaaggg caacaggaag ggcatttagga 60
 cttgaaatcg gagacgcacg caggggaggg agtcagtgtc ggaacctggt aggcctggg 120
 agaactccgg ctttcgtct gcgtgagctg gagaagagcc gaaggttct gcgcacagca 180
 cggacctgcg tgccctcagct ttaaggaaat caccgtggcc gccgctgtga acgcagagaa 240
 gggcgcgagc gtgggagcag gaacccaagg cggtgggaaa cggtggggct ttctgagtgt 300
 attggaaagt agagcccaca gatctgctgc agaccagaaa ggggcgcgag aaagagcgg 360
 cagaggcaga cgccggggct ggcggcgatg gagcagcagt cggaggacgc ggaaggcctg 420
 cgagagtcgc ccgcggccca ggcggccct tcgggcccac cttgcgggt gatgttgtgc 480
 acgttaggggc acgtgttgca ggcgaagcgg tggcagcgtt gtccctcctc cacgatcagc 540
 ccgttcccgcc agccggggca gaacagcagc atggtctcga actccgcagg ctccaaactcc 600
 cggcagctcc cactgcccgt cagcgccgtat ggcggccccc cctcgagctc acattggtcc 660
 tggcagcctt cccggcacac caaccaacca atagacaggg cgattctgcg ctcccgccct 720
 gctgcagggct gtctcgact tgtcattggt cactgcagcc gccccacccc cccggcgcc 780
 cagtggctgg gcccgcgc tggggcggggc cgcagttct ggcgtgcgc gttggcctc 840
 cctagtgcgg gctggcagtg cgggcagagc cggcgtgaga ggggcggccc tggaggagac 900
 ggaggcggcg ggtggggcccg aggcgcagaaga ggaagatgag gacgaagaag aggccgtgcc 960
 gcactcccgag gccatggacg tggcctcaggaa gggtctggct atggtgggtgc aggacccgt 1020
 gctctgcgtatcc ctgcggatcc aggttactct ggaagaagtc aactccaaatagccctaga 1080

atacggccag gcaatgacgg tccgagtgtg caagatggat ggagaagtaa tgcccgtggt	1140
tgttagtgcag agtgcacag tcctggacct gaagaaggcc atccagagat acgtcagct	1200
caagcaggag cgtgaagggg gcattcagca catcagctgg tcctacgtgt ggaggacgta	1260
ccatctgacc tctgcaggag agaaactcac ggaagacaga aagaagctcc gagactacgg	1320
catccggaat cgagacgagg tttccttcat caaaaagctg aggcaaaagt gaggccctccag	1380
acaggacaac cctcttcatc actggtggtc gagcttttc ccagcaggaa tgggtcctcg	1440
aatcatcgta cctctttcac agaaaggacg ttgtggtggc ctcaccccaag gcatgccccaa	1500
caggaactgt cagcattaaa cctgggggcc ctcaggacta ggacagggtg agccagtgtc	1560
ccctccttc atgtacttgg cctgagactg acctctccct aggtccaaat gccctagtca	1620
catggagaca cggctggcac tgtaataaa ctgttggttt agttgaagga caaaaaaaaaa	1680
ggggcgggtg aagttactct ggggcgagta ggaccgttt ggaaaggcca tgtgggatta	1740
agagaagggg ggtaaagtgc gaaaagcatg gtttggagag attgggggaa gagagcgaga	1800
ggagggggaaa ggtgagaagg gggaggtgtta taagagagg	1839

<210> 55
<211> 2586
<212> DNA
<213> Homo sapien

<400> 55	
ggcacgaggg agagatgtc caggaggcgg tggatgtcg gcaaaggca acaggaagg	60
cattaggact tggaaatcgga gacgcacgca ggggagggag tcagtgtcg aacctggtag	120
gccctggag aactccggct ttctgtctgc gtgagctgga gaagagccga aggtttctgc	180
gcacagcacg gacctgcgtg ctcagcttt aaggaaatca ccgtggccgc cgctgtgaac	240
gcagagaagg gcgcgagcgt gggagcagga acccaaggcg gtgggaaacg gtggggcttt	300
ctgagtgtat tgaaaagtag agcccacaga tctgctgcag accagaaagg ggccgcgagaa	360
agagcggaca gaggcagacg ccggggctgg cggcgatgga gcagcagtcg gaggacgcgg	420
aaggcctgcg agagtcgccc gcggcccaagc gccggcccttc gggtcccacc ttgcgggtga	480
tgttgtgcac gtggggcac gtgttgagg cgaagcggtg gcagcgttgt ccctcctcca	540
cgatcagccc gttcccgcaag ccggggcaga acagcagcat ggtctcgaac tccgcaggct	600
ccaactcccc gcagctccca ctgcccgtca gcggcgatgc gccggccgc tcgagctcac	660
attggtcctg gcagccttcc cgccacacca accaaccaat agacagggcg attctgcgt	720
cccgccctg ctgcaggctg tctcgactt gtcatggtc actgcagccg cccacccccc	780
cccgccgcgc cagtggctgg gcggcctcgc tggggcgggc cgcatgttgcgc	840

gcttggcctc	cctagtgcgg	gctggcagtg	cgggcagagc	ccggctgaga	ggggcgcccc	900
tggaggagac	ggaggcggcg	ggtgggccc	aggcgcaaga	ggaagatgag	gacgaagaag	960
aggcgctgcc	gcactccgag	gccatggacg	tgttccagga	gggtctggct	atggtgtgc	1020
aggaccgc	gctctgcgat	ctgcccgcatt	aggttactct	ggaagaagtc	aactccaaa	1080
tagccctaga	atacggccag	gcaatgacgg	tccgagtgtg	caagatggat	ggagaagtaa	1140
tgcccggtt	tgttgtgcag	agtgcacag	tcctggacct	gaagaaggcc	atccagagat	1200
acgtgcagct	caagcaggag	cgtgaagggg	gcattcagca	catcagctgg	taagtggAAC	1260
aacattccct	tcattatagc	ccttcgtggg	gctagtgc	ttcttggcac	tgtcaccagg	1320
caccacctgg	aaacagctct	cagctctgca	tgagtacagc	accactgaag	tgtgagctc	1380
cctgtcacaa	gagtgtat	ctccctgtca	cagacagtgc	gggtcggttct	gtgcctggga	1440
ctcctgcctc	ggccatcccc	aacattctgc	tcttcatcg	gcatcacccc	atccgagctg	1500
ctgggtatct	tcacttgggg	acactgtcgg	gaatttccag	tgtgtctgga	agtggctcc	1560
ctagtttgg	atggtacacc	tgttagggct	cccatcccc	tctcacctgg	gtgctgtcag	1620
ccctcactct	cctattggat	caactatcct	gttcaactgag	tctcaacact	gtgcctgtt	1680
gcattagcaa	ggtttgtttg	gccaagccgc	cccagacagc	cctctgagaa	cagacccctcc	1740
ttgttagctgc	ctcagaccca	atctgcacat	tgtacagaac	agcccaggta	gggaggacag	1800
ctgccccagg	tcccatagga	ctgcatgc	caagcccacg	tcatgcagag	ccactcagct	1860
caccctgctc	agggcacgtg	gtttacctgc	attccccctct	tgcaggtcct	acgtgtggag	1920
gacgtaccat	ctgacacttg	caggagagaa	actcacggaa	gacagaaaga	agctccgaga	1980
ctacggcata	cggaatcgag	acgaggttgc	cttcataaaa	aagctgaggc	aaaagtgagc	2040
ctccagacag	gacaaccctc	ttcatcaactg	gtggctgagc	tttttccag	caggaatggg	2100
tcctcgaatc	atcgtgcctc	tttcacagaa	aggacgttgt	ggtggcctca	ccccaggcat	2160
gccccacagg	aactgtcagc	ataaacctgg	gggcctcag	gactaggaca	gggtgagcca	2220
gtgctccctc	cttcatgtt	cttggcctga	gactgacctc	tcccttaggtc	caaatgcct	2280
agtcacatgg	cagacccacg	gcctggccca	ctgtataaaaa	taaacctgtt	tgcttcttat	2340
cttagtttga	aaagtagaaa	gccacagtaa	cctgggtagc	aaagactgag	attgccccat	2400
cacagagggt	agttaaagggg	agagaattgg	tacaggcgag	tcctatagtc	caagatggcg	2460
ccacaccacc	aaagccttga	ggccacacca	ctccccaaac	cacacaactg	tgttaccatg	2520
atctccacag	caaggaggaa	ataaaagcag	agcggcttta	gggtttgcat	cctggagctc	2580
acagtg						2586

<210> 56
<211> 2566
<212> DNA
<213> Homo sapien

<400> 56
ggcacgaggg agagatcgtc caggaggcgg tggatgcg gcaaaggca acaggaaggg 60
cattaggact tgaaatcgga gacgcacgca ggggagggag tcagtgtcg aacctggtag 120
gccctggag aactccggct ttctgtctgc gtgagctgga gaagagccga aggtttctgc 180
gcacagcacg gacctgcgtg cctcagctt aaggaaatca ccgtggccgc cgctgtgaac 240
gcagagaagg gcgcgagcgt gggagcagga acccaaggcg gtgggaaacg gtggggctt 300
ctgagtgtat tggaaagtag agcccacaga tctgctgcag accagaaagg ggccgcgagaa 360
agagcggaca gaggcagacg ccggggctgg cggcgatgga gcagcagtcg gaggacgcgg 420
aaggcctgcg agagtcgccc gcggcccagc gccggcccttc gggtcccacc ttgcgggtga 480
tgttgtcac gtaggggcac gtgttgcaagg cgaagcggtg gcagcgttgt ccctccctca 540
cgatcagccc gttcccgcaag ccggggcaga acagcagcat ggtctcgaac tccgcaggct 600
ccaactcccc gcagctccca ctgcccctca gcgcgcgtgc gccgcggcc tcgagctcac 660
attggtcctg gcagccttcc cgccacacca accaaccaat agacagggcg attctgcgt 720
cccgccctg ctgcaggctg tctgcactt gtcatggtc actgcagccg ccccaccccc 780
cccgccgcgc cagtggctgg gcggcctcgc tggggcgggc cgcagttcct gcgcgtgcgc 840
gcttggcctc cctagtgcgg gctggcagtg cggcagagc cccgctgaga ggggcggccc 900
tggaggagac ggaggcggcg ggtggcccg aggcgcaga ggaagatgag gacgaagaag 960
aggcgctgcc gcactcccgag gccatggacg tggatggca gggtctggct atgggtggc 1020
aggaccgcgt gctctgcgtat ctgcgcgtcc aggttactct ggaagaagtc aactccaaa 1080
tagccctaga atacggccag gcaatgcgg tccgagtgta caagatggat ggagaagtaa 1140
tgcgtaaatg ctaccctccct cccttcaggat tatgtggtcc aggctttcac agcaggaaga 1200
cctaacagtg ctggtcagcc tgctcagaaa ctcacaggcc atgcccagggt gtactgggc 1260
aaccacaaac ctgcctgtg cacagaggtg ttggatccctt tccgcgtccatc ggaggctgt 1320
gctttgggtt ctcaccatgg atcttctccc atctgtgtcc gtggatggcag ccgtgggtgt 1380
agtgcagagt gccacagtcc tggacctgaa gaaggccatc cagagatacg tgcagctcaa 1440
gcaggagcgt gaagggggca ttcaagcacat cagctggtaa gtggaaacaac atcccttca 1500
ttatagccct tcgtgggct atgtcccttc ttggcactgt caccaggcac cacctgaaaa 1560
cagctctcag ctctgcgtatgatgatgatgatgatgatgatgatgatgatgatgatgatgatg 1620

71

tgatgagctc	cctgtcacag	acagtgcggg	tcgttctgtg	cctgggactc	ctgcctcgcc	1680
catccccaaac	attctgtct	tccatcgca	tcaccccatc	cgagctgctg	ggtatcttca	1740
cttggggaca	ctgtcgggaa	tttccagtgt	gtctggaagt	ggcctcccta	gtttggatg	1800
gtacacctgt	aggggctccc	atcccctct	cacctgggtg	ctgtcagccc	tcactctcct	1860
attggatcaa	ctatcctgtt	caactgagtct	caacactgtc	gcctgttgca	ttagcaaggt	1920
ttgtttggcc	aagccgcccc	agacagccct	ctgagaacag	agcctccttg	tagtgcctc	1980
agacccaatc	tgcacattgt	acagaacagc	ccaggttaggg	aggacagctg	ccccagggtcc	2040
cataggactg	catgcctcaa	gcccacgtca	tgcagagcca	ctcagctcac	cctgctcagg	2100
gcacgtggtt	tacctgcatt	ccccctttgc	aggccttacg	tgtggaggac	gtaccatctg	2160
acctctgcag	gagagaaaact	cacggaagac	agaaaagaagc	tccgagacta	cggcatccgg	2220
aatcgagacg	aggtttcctt	catcaaaaag	ctgaggcaaa	agtgagcctc	cagacaggac	2280
aacctcttc	atcaactggtg	gctgagcttt	ttcccagcag	gaatgggtcc	tcgaatcatc	2340
gtgcctcttt	cacagaaagg	acgttgtggt	ggcctcaccc	caggcatgcc	caacaggaac	2400
tgtcagcata	aacctggggg	ccctcaggac	taggacaggg	tgagccagtg	ctccctcctt	2460
tcatgtactt	ggcctgagac	tgacctctcc	ctaggtccaa	atgccctagt	cacatggcag	2520
acccacggcc	tggcccactg	tataaaataa	acctgtttgc	ttctta		2566

<210> 57
<211> 2817
<212> DNA
<213> Homo sapien

<400> 57						
gcccactttgc	gtcacgtcc	actgccactc	tcacggaaac	tcctacaaga	acggcacaca	60
cgttcgctcc	ctcagcatttgc	caaacacgct	ccccccaaaa	ccacaaacgc	gcccccacac	120
actcgcttac	tttccctcac	aaagatgccc	gcttacacgg	ccacagcagg	cacgctcaga	180
gacacgcagt	tacacacacaca	catcgctgtt	cacaacccca	catacaatca	aaaaacaaaa	240
cacgaaacgt	tccccctgggc	actaaatcct	cacgttaacg	tacacacacaca	aacacacgccc	300
ctcctctccc	acttccttctt	ccataccctt	tcctcgagggc	ccccccaccc	tgattttcgg	360
caccccccagt	cccaatcata	attggcgccc	gcgcagccct	ctttggacac	acacgcgcgg	420
cccacgcacg	cgctccccctc	ccggcgaggc	ggggggcggt	tcggccggga	gccggcgagg	480
cccgcttcgg	attccaggttgc	tggcagtgtac	tccgcgcgtcc	acgttttgca	ggctgccacc	540
gtgtcggagg	cgaggcgagg	aagggagctg	gaataacaaa	gggaggtaat	gggtagatt	600
ggataaccttct	ggggtcttgg	aagaagctat	gacttattta	ctgtctacta	tgtggccctg	660

72

acattctcca	gcttcatgg	tgttcctgca	atccaaagtgc	cgccttatctc	tctacaggc	720
tggaagaaaag	ccataacttct	gactccagga	agtgcattgg	tggagctgga	ccggcattctg	780
agctgtctgt	ggcggtccacc	agatgaccgg	ctggcctctt	aattgggctt	cctgcattcc	840
tgcctgtgct	cggagctccc	ttctccactt	gcctgctgga	ggaggtttac	aaacaaaatc	900
agatcatggc	actccccctgg	ggccctgcat	cctgtggcca	gcaggtcttg	gaaacgcccc	960
tgtaccccag	aaacttcctg	atcagtcctt	ggtcagtgtt	tatcatgctt	cctgtgacat	1020
gacatcaact	tttcagtgac	ttccactggt	tttggcccaag	ctcgaagagc	ctgacacacct	1080
tgcagggaat	tctaccaggg	agaagagaaa	gtagtcactg	cttccagttt	tttaaggag	1140
ttggccaaga	gccagcacca	ggcatcagag	gaaggcagct	gactgcactc	ggccacatcc	1200
caaaaagtgcc	tggaagggga	gggaggaagc	aggcgcttca	gaaggcacta	ctgtgtgtca	1260
ggactcatgc	taggaacgct	gtaacaggg	gcagtgtatgg	agtccctttag	ggagctcagg	1320
agggggaaatt	gcaactgaaga	tttgccttcat	cgtgaagcaa	gagcactgga	acttacactc	1380
caccatcagg	ccctgtcaca	ggagaacaaa	gaaggaaggc	agaggagatc	atgctccggc	1440
cagcaggaa	tctccatttt	tttcagccctc	ataccttgg	aagtacaaag	gagtgagagc	1500
tgggactacc	agccagaggg	ttcatggagg	tagtggagg	ggaagatgg	ttctgcatgg	1560
agccactcca	gggacttttc	ttctgtctca	cagcctgaca	atcacctcta	gctgttctca	1620
gtccattct	catcaatgac	ctcgctctt	gattttgtac	aatatctagg	tgaccccttc	1680
ttctttccat	tttcaaacgc	ttgatcattt	atgccttctt	ttctttcttc	tctccctctg	1740
ggaaaagaga	aggaccaaca	tcccccttcc	cattttctt	gtgctgtccc	gaccctctg	1800
agacctggct	ctagcagtaa	gtccctctct	cttctgcagc	accagcctct	gccccatccct	1860
gccctctgca	cacaaagctc	tcacattgtc	tgcactttaa	gaactttgt	ggcaaaccct	1920
ggcaaattct	gccacttgg	tccacttcttc	tcaaattggct	ttgaggtcat	cattgaccta	1980
tcccatattga	ccctgtcttg	tatctgtatgt	tgtacccttt	ctccttgatc	ccactctggc	2040
aactctccca	ctcccccttcc	ttggatcaac	tgtcttgctc	ctcagttctg	atcaatgcct	2100
ccttctctct	ctttagcctt	ggcccttttt	ggccaggaa	ataaaatgtga	aatacataaa	2160
gcttcatttt	atcttatgt	tatagctatg	ttgatgttac	ccaagaccta	aaagtacacac	2220
tctctaatta	cacttccctc	ctgtgttctg	ggtccttctt	gaccagtctt	tactctggct	2280
ccccctctgg	atgtctccca	gacttctaat	tagaattact	gtttcttata	tcaagctaat	2340
tatctttccc	accaacctcc	tccatcctt	ggtgaaagca	tcacttgggt	gaaaaagtca	2400
gaaatctggg	tttcattttt	gcggtaggtt	gaataatgcc	cccaggcccc	aggtttgac	2460
atccgaatcc	ctgggacttt	gcagatgtt	ttcagtttag	gattttgaga	tggggagatt	2520

atctgggaga	gcctgatgtc	atcataaggt	tcttataaga	gggaggcagg	agggttagag	2580
tgagtagtaa	gagatgcaac	agtggaaagca	agaggttggg	gtgatgtggc	cacaagccca	2640
ggagtgctga	cagacatcag	aagctggaag	ggacaaggaa	tggttctcc	tctggagcct	2700
ccagaaaagaa	ccagccctgc	tgacaccttg	attttagcct	tggaagactc	attttggact	2760
tctgacacctg	aacgttgtaa	gagaataaat	ttacatattt	taaactggaa	tgttat	2817

<210> 58
<211> 1530
<212> DNA
<213> Homo sapien

<400> 58						
atctagctct	gcatgccacc	cagggagctc	aggaggggga	attgcactga	agatttgc	60
catcgtaag	caagagcact	ggaacttaca	ctccaccatc	aggccctgtc	acaggagaac	120
aaagaaggaa	ggcagaggag	atcatgtcc	ggccagcagg	aatctccat	tttttcagc	180
ctcatacctt	ggaaagtaca	aaggagttag	agctggact	accagccaga	gggttcatgg	240
aggttagtggg	aggggaagat	gggttctgca	tggagccact	ccagggactt	ttcttctgtc	300
tcacagcctg	acaatcacct	ctagctgttc	tcagtcctat	tctcatcaat	gacctcgct	360
cctgattttg	tacaatatct	aggtgacett	ctcttcttcc	catcttcaaa	cgcttgcata	420
ttgatgctcc	ctttctctt	ctctctctc	ctggggaaag	agaaggacca	acatccccct	480
ttcccttcct	cctgtgctgt	cccgaccctt	ctgagacctg	gctctagcag	taagtcctc	540
tctctctgc	agcaccagcc	tctgcccattc	cctgcccctt	gcacacaaag	ctctcacatt	600
gtctgcactt	taagaacttt	tgtggcaaacc	cctggcaaatt	cctgccactt	gtttccactc	660
ttctcaaatg	gctttgaggt	catcattgac	ctatcccatt	tgaccctgtc	ctgtatctga	720
tgttgaccc	tttctccttg	atcccactct	ggcaactctc	ccactccccct	ttcttggatc	780
aactgtcttg	ctccctcagtc	ttgatcaatg	cctccttctc	tctctttagc	cttggccctt	840
tttggccagg	gaaataaaatg	tgaaatacat	aaagcttcat	tttatcttat	gattatagct	900
atgttgatgt	tacccaagac	ctaaaagtca	cactctctaa	ttacacttcc	ctcctgtgtt	960
ctgggtcctc	ttggaccagt	ccttactctg	gctccctcc	tggatgtctc	ccagacttct	1020
aattagaatt	actgtttcct	atatcaagct	aattatcttt	cccaccaacc	tcctccatcc	1080
ttgggtgaaa	gcatcacttg	gttgcaaaag	tcagaaatct	gggtttcatt	cttgcggtag	1140
gtagaataat	gcccccaggc	cccaggtttt	gacatccaa	tccctggac	tttgcagatg	1200
tgattcagtt	taggattttg	agatggggag	attatctggg	agagcctgat	gtcatcataa	1260
ggttcttata	agagggaggc	aggagggta	gagttagtag	taagagatgc	aacagtggaa	1320

gcaagaggtt	gggtgatgt	ggccacaagg	ccaggagtgc	tgacagacat	cagaagctgg	1380
aaggacaag	aatggtttc	tcctctggag	cctccagaaa	gaaccagccc	tgctgacacc	1440
ttgattttag	ccttggaga	ctcattttgg	acttctgacc	ttgaacgtt	taagagaata	1500
aatttacata	ttttaaactg	aatgtttat				1530

<210> 59
<211> 3490
<212> DNA
<213> Homo sapien

<400> 59						
taagcctcat	agtctaagaa	agccctcaag	caaggctaac	attttggta	tctgcgagaa	60
gattgagcac	tcggtgtcct	tgctccttgc	agcttcgcag	catcttctgg	agcagcatga	120
gcttcact	ctgactcata	agtctccac	cctcataaggc	cccactgggg	agtttggggg	180
cctctattgc	catgtgcctg	gaattattat	atgctcatca	ctttatgaas	aayaaaattt	240
gtcttkcctg	ccttaaagtt	acattcggtc	ttccgctcaa	atcctgatct	ggtcattaa	300
agagtgttcg	cagacaaagt	ttctgaaaga	ttagagaaga	atccccccca	agattgcccc	360
aacactgaac	tacagacaaa	cactatttt	tttaaataag	gagacagctt	tctaaaagta	420
tacattctct	aataaaaata	gtttattatt	ttgaatgatt	taatggttt	ctacacaatt	480
tacatcacaa	catgtaaattt	tttagcagtaa	catctgattc	taacagcaca	tcatgctatt	540
ccttcatag	agccttcaga	gattcaatgc	taaacaattt	tccttagttt	gcatcaaggc	600
actgatca	tttagaggctt	ttaagaaattt	attnaaagat	gcaaattgcct	ctgagtgaag	660
tgtactatcc	catcaactgaa	gcccacagga	acaagtccct	caatttaaa	aaggctcgat	720
ggaaaaattt	ctcaatcctg	aaatccccta	gggaagggg	caggagaaag	tgccatggtt	780
gatatttaag	aactccacag	ctcttaaaaa	taagcactta	tccctaacat	gcaataactgc	840
agatgcaagt	taaacttac	tgttaacagc	tgcctgctgt	tttctgctcc	cagatgaaat	900
gaagcaactc	ttctgataac	gaagagatac	ctgtctgagg	caaacgaaac	attggcacac	960
agcacagcct	cctcaatcca	cttgatccca	actcatctt	catttatttc	ggcttctttt	1020
atccaggat	taatgttagt	taacattttc	atttcttttc	gcttttattt	tgctttgtt	1080
aaagcagtat	tttgagatgg	acattgcctc	ttcattgtat	ttctcatcaa	ttcatttattt	1140
ttgtggttat	agcttgacaa	gcaattaact	ttaaaatgg	agattccgta	actttaaattt	1200
ggtagcttcc	atttgcttaa	aatttttgg	catatgcaga	taatgttctc	atcagtagta	1260
agaatctcag	ggttatgctt	attcccaat	ggaggtatga	catataatct	tttctgcctt	1320
tacttatcaa	ttcaccaagg	agctgtttc	tctgcata	ggccatcata	ctgccaggct	1380

ggttatgact cagaagatgt tatctgaaaa aagtctata tagaaaaaaaaa artktccccct 1440
ccctcatcaa caaaagccca ccctctaaga gacattcaag ctgaactatc acaattctta 1500
atcagttaca atttacaaac agataagtt aaaataaaaca atttacaaaa ttttgaagc 1560
ataccttaac atcttgaaaa gcagttaaac aatggaaaag tatttctcct acactaaaaa 1620
aaaacttgct tacacacaac tgaaaataga atcttacttg ataatacaaa agctaccatc 1680
agaagaaaatc ctttcaggat cattaagcca cttcccccgc tctgcagttt ctatagtagt 1740
tttaaattat tattaaatca cctgaaaaaa attccaaaag agaaccacac actaccatat 1800
ccaaacaact ttgcatttc ccataattgt agttaatgtc agcccagtag gccagaccaa 1860
cccccagttc aatactttcc ttccccaaaa gctctatact ttgaaggaaa acagatacag 1920
tatcaaatta tgacactttc cttgccccaa ttaatgcact ggtacacccca gtggctcata 1980
tttaacttcc cccagcttcc caattcaaac tggggggaaa aaaactaaat cattggaggt 2040
tacttgccaa cttggaagtt gatatttctt tacttttcc attctaagac tttaagttct 2100
ctggcatgag tttatctgca atcataaact aaacaattac ctaaaccac cccaccaatc 2160
ccaaccgtaa caggccactg ccaactaatt gccaatattt gcccctcccc tttaataaaaa 2220
cttttaagaa gtcacattat tggaaaactt aacttcaaca tttggcctac tcaagctctt 2280
ctgaagttct cctgagatga ctgaatatga accaaagctg cactgtgctg tactttcag 2340
cttcaactgg gaatacttcc ccaaggataa aagcagctcc agtccctgaa ggtgttcgtg 2400
ccaaacagcac agcggtacat tcaccaaatac gcactggctc ctggactctt ttcctatctt 2460
caccacgaac tgctgcttgc tcgcttgc ctcagtccta gcttcatcaa acactggttc 2520
ctggaatcct gtctgctgct gtcttcctag attcactgaa tccacttctg tgttagcacct 2580
gggtcagctg tcaattaatg cttagtccta ggattaaaa aataatctta actcaaagtc 2640
caatgcaaaa acattnaagtt ggttaattact cttgatctt aattacttcc gttacgaaag 2700
tccttcacat ttttcaaaact aagctactat atttaaggcc ttccaaattt ttcttaactct 2760
tccaaaagcc ttctgcctta gttttttta aattacacca gtccttttag tagcttttg 2820
atgtgattt taaccaactt ccccttctag cttcaagttat tcttctaaat tggttctgg 2880
ctacgtaaac accctcatct tctcaagctt taccttctaa cttctgcacc accagaaatt 2940
aaatttgatgg gctttaaaaa taaattgggtt accaataatt tcctcatttt ttcagtgcta 3000
ttttatccaa tttttggctt tatatttttc tatcttctat acttctccaa tacttgctt 3060
agcttggttt tcattttcta tctgaaaactc ttgacaatat ttctcattttc tatcttgctt 3120
ctatcttcca attttcttct aagtttgatc attttgcct tagcttttg ttctcttagt 3180

76

tgtctttttt cttctgcttc ctactttca ggtttaaatt tatctttttt cttctaaaag	3240
tatgtttta tcttctaatt tccctatctt ctctattctt ttcttcgcct tcccgtactt	3300
ctgtcttcca gtttccact tcaaacttct atcttctcca aattgttca tcctaccact	3360
cccaattaat ctttccattt tcgtctgcgt ttagtaaatg cgtaactag gctttaatg	3420
acgcaattct ccctgcgtca tggatttcaa ggtctttaa tcaccttcgg tttaatctct	3480
ttttaaaaga	3490

<210> 60
<211> 2238
<212> DNA
<213> Homo sapien

<400> 60	
taaggcctcat agtctaagaa agccctcaag caaggctaac attttggtaa tctgcgagaa	60
gattgagcac tcggtgtcct tgctccttcc agcttcgcag catcttctgg agcagcatga	120
gcttctact ctgactcata agtctccac cctcataagc cccactgggg agtttgggg	180
cctctattgc catgtgcctg gaattattat atgctcatca ctttatgaas aayaaaattt	240
gtcttkcctg ccttaaagt acattcggtt ttccgctcaa atcctgatct ggtccattaa	300
agagtgttcg cagacaaaagt ttctgaaaga ttagagaaga atccccccca agattgcccc	360
aacactgaac tacagacaaa cactattta tttaaaataag gagacagctt tctaaaagta	420
tacattctct aataaaaata gtttattatt ttgaatgatt taatggttt ctacacaatt	480
tacatcacaa catgtaaatt ttagcagtaa catctgattc taacagcaca tcatgtatt	540
ccttcatag agccttcaga gattcaatgc taaacaaatt tccttagtg gcatcaaggc	600
actgatcaact ttagaggcatt ttaagaaatt atttaaagat gcaaattgcct ctgagtgaag	660
tgtactatcc catcaactgaa gcccacagga acaagtccta caattttaaa aaggctcgat	720
ggaaaaattt ctcacatctg aaatcccta gggaaagggtt caggagaaag tgccatggtt	780
gatatttaag aactccacag ctcttaaaaa taagcactta tccctaacat gcaatactgc	840
agatgaagt taaaacttac tggtaacagc tgctgtgtt tttctgctcc cagatgaaat	900
gaagcaactc ttctgataac gaagagatac ctgtctgagg caaacgaaac attggcacac	960
agcacagcct cctcaatcca cttgatccca actcatctct catttatttc ggcttctttt	1020
attccaggat taatgttagtg taacatttcc atttcttttc gcttttatttc tgctttgtt	1080
aaagcagtat tttgagatgg acattgcctc ttcatgttat ttctcatcaa ttcatttattt	1140
ttgtggttat agcttgacaa gcaattaact taaaatggt agattccgta actttaaattt	1200
ggtagcttc atttgcttaa aatttttgg catatgcaga taatgttctc atcagtagta	1260

agaatctcag ggttatgctt attcccaat ggaggtatga catataatct tttctgcctt	1320
tacttatcaa ttcaccaagg agctgtttc tctgcata ggcatacgat ctggcaggct	1380
ggttatgact cagaagatgt tatctgaaaa aagtctatag aaaaaaaaaa artkccccct	1440
ccctcatcaa caaaagccca ccctctaaga gacattcaag ctgaactatc acaattctta	1500
atcagttaca atttacaaac agataagttt aaaataaaaca atttacaaaa ttttgaagc	1560
ataccttaac atcttgaaaa gcagttaaac aatggaaaag tatttctcct acactaaaaa	1620
aaaacttgct tacacacaac tgaaaataga atcttacttg ataatacaaa agtaccatc	1680
agaagaaatc ctttcaggat cattaagcca cttcccccgc tctgcagttt ctatagtagt	1740
tttaaattat tattaaatca cctgaaaaaa attccaaaag agaaccacac actaccatat	1800
ccaaacaact ttgcatttc ccataattgt agttaatgtc agcccagtag gccagaccaa	1860
cccccagttc aatactttcc ttccccaaaa gctctatact ttgaaggaaa acagatacag	1920
tatcaaatta tgacactttc cttgccccaa ttaatgcact ggtacaccca gtggctcata	1980
tttaacttcc cccagcttcc caattcaaac tggggggaaa aaaactaaat cattgggagtt	2040
tacttgccaa ctggaaagtt gatatttctt tacttttcc attctaagac tttaagttct	2100
ctggcatgag ttatctgca atcataaaact aaacaattac ctaaaccac cccaccaatc	2160
ccaaccgtaa caggccactg ccaactaattt gccaatattt ggagggatga gcataaggag	2220
ggatgagcat atgagggt	2238

<210> 61
<211> 2226
<212> DNA
<213> Homo sapien

<400> 61	
taagcctcat agtctaagaa agccctcaag caaggctaac attttgtca tctgcgagaa	60
gattgagcac tcgggtcct tgctccttc agcttcgcag catcttctgg agcagcatga	120
gcttctcaact ctgactcata agtctccac cctcataagc cccactgggg agtttgggg	180
cctctattgc catgtgcctg gaattattat atgctcatca ctatgaas aayaaaattt	240
gtcttkcctg cttttaagtt acattcggttc ttccgctcaa atcctgatct ggtccattaa	300
agagtgttcg cagacaaagt ttctgaaaga ttagagaaga atccccccca agattgcccc	360
aacactgaac tacagacaaa cactattta tttaaataag gagacagctt tctaaaagta	420
tacattctct aataaaaata gtttattatt ttgaatgatt taatggttt ctacacaatt	480
tacatcacaa catgtaaattt ttagcagtaa catctgattc taacagcaca tcatgtatt	540
ccttcatag agccttcaga gattcaatgc taaacaaatt tccttagtt gcatcaaggc	600

78

actgatcact ttagaggc tt	aagaatt attaaagat gcaa	atgcct ctgagtgaag	660
tgtactatcc catca	cgtaa gcccacagga acaagtccta caat	ttaaa aaggctcgat	720
ggaaaaattt ctc	aatccccctg aaatccccct gggaggggt caggagaaag tgccatgg	tggtt	780
gatatttaag aactccacag ct	ctttaaaaaa taagcactta tccctaacat gca	ataactgc	840
agatgcaagt taaacttac	tgttaacagc tgccgtgt tttctgctcc cagatgaaat	900	
gaagcaactc tt	ctctgataac gaagagatac ctgtctgagg caaacgaaac attggcacac	960	
agcacagcct cctcaatcca	cttgatccc actcatctc catttat	ttc ggcttctttt	1020
attccaggat taatgtat	gta acattttc atttcttgc gctttat	tc tgctttgt	1080
aaagcagtat tt	tgagatgg acattgcctc ttcatgtat ttctcatcaa ttcattat	ttttt 1140	
ttgtggttat agcttgacaa	gcaattaact taaaatggt agattccgta actttaaatt	1200	
ggtagcttc atttgcttaa	aatttttgg cata	tgca gaaatgtc atcagtagta	1260
agaatctcag gttatgctt	attcccaat ggaggtatga catataatct tttctgc	ctt 1320	
tacttatcaa tt	caccaagg agctgtttc tctgcata ggcata ctgcaggct	1380	
ggttatgact cagaagatgt	tatctgaaaaa aagtctatag aaaaaaaaaa artktccc	1440	
ccctcatcaa caaaagccca ccctctaaga gacattcaag ctgaactatc acaattctt	a 1500		
atcagttaca atttacaaac agataagtt aaaataaaca atttacaaa tt	tttgaagc 1560		
atcaccta ac	ttttgttt gcagttaaac aatggaaaag tatttctc	actacaaa 1620	
aaaacttgct tacacacaac tgaaaataga atcttacttg ataatacaaa agtaccatc	1680		
agaagaaatc cttcaggat cattaagcca cttccttgc tctgcagtt ctat	agtagt 1740		
tttaaattat tattaaatca cctgaaaaaa attccaaaag agaaccacac actaccat	1800		
ccaaacaact ttgcatttc ccataatgt agttaatgtc agcccagtag gcca	gacc 1860		
cccccagttc aatactttcc ttccccaaaa gctctatact ttgaaggaaa acagatac	1920		
tatcaaatta tgacacttcc cttgccccaa ttaatgcact ggtacacccaa gtggc	cata 1980		
tttaacttcc cccagttcc caattcaaac tggggggaaa aaaactaaat cattggg	aggt 2040		
tacttgccaa cttggaagtt gatatttctt tacttttcc attctaagac tt	taagttct 2100		
ctggcatgag ttatctgca atcataaact aaacaattac ctaaacccac cccaccaatc	2160		
ccaaccgtaa cagggcactg ccaactaatt gccaatattt tacctgc	ccg cgaccacgct 2220		
aaggc		2226	

<210> 62
<211> 981
<212> DNA
<213> Homo sapien

<400> 62
tgctttgtg tctacttcct tgtgcctmc ggagtcgagc tctgtcagtg catgattctt 60
gccaatcgct aaacgttagga ctcgaggaag gccattggca attctgtcaa gaagacagtg 120
caagtgccta taaaaacgac agtttagggg gaaaacaaac caatacccg aaagctgaga 180
ggccagctt ttaatcgta tggtttatg taaaataaaa cagcacgtgg aaggattgt 240
aagcgcttgg tggctgctcg agccccaga aaggtgctt gttcttccac ctctgccact 300
aattcgacat cagtttcatc gaggaaagct gaaaataaaat atgcaggagg gaaccccg 360
tgcgtgcgcc caactccaa gtggaaaaa ggaattggag aattctttag gttgtcccct 420
aaagattctg aaaaagagaa tcagattctt gaagaggcag gaagcagtgg cttagaaaa 480
gcaaagagaa aagcatgtcc tttgcaacct gatcacacaa atgataaaaa agaatagaac 540
tttctcattc atctttgaat aacgtctcct tggttaccct ggtattctag aatgtaaatt 600
tacataaaatg tgtttgtcc aattagcttt gttgaacagg catttaatta aaaaatttag 660
gtttaaattt agatgttcaa aagtagttgt gaaatttgag aatttgaag actaattatg 720
gtaacttagc ttagtattca atataatgca ttgtttgggt tctttacca aattaagtgt 780
ctagttcttgc ctaaaatcaa gtcattgcat tgggttctaa ttacaagtat gttgtatgg 840
agatttgctt agattgttgc actgctgcca ttttattgg tggttgatta ttggaaatgg 900
gccatattgt cactccttct acttgcttta aaaagcagag ttagatttt gcacattaaa 960
aaattcagta ttaattaaac a 981

<210> 63
<211> 706
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (34)..(34)
<223> n=a, c, g or t

<400> 63
cccccccccc cgccctactta tctataaggg ccantggta tcctagatgc tgctcgagcg 60
gcgcagtgtg atggattggc cgccggccgag gtaccagatt ataatgccag aatataatgt 120
gcaggcaatc gtggatgtct ctgacaaagt gtgtctcaa aataatatac ttttacatta 180
aagaaattta atgtttctct ggagttgggg ctcttggctt tcagagttt gttatcagt 240
gttgattcta gatgatcaac ataatggacc actcctgaat gagacttaat ttgttcttc 300
aaatttactg tcttaaatca gtttattaaa tctgaatttt aaaacatgct gtttatgaca 360

80

caatgacaca ttgttgcac caattaagtg ttgaaaata tcttgcatc atagaacaga	420
aatatataaa aatatatgtt gaatgttaac aggtatttc acaggttgt ttcttgatag	480
ttactcagac actagggaaa ggtaaataca agtgaacaaa ataagcaact aaatgagacc	540
taataattgg cttcgattt taaatatttgc ttcttataaa cttgtcaat aaaaataaat	600
ctaaatcaga aaaaaaaaca acacaaaaaa aaaagggtgg gggaaaccag ggccaaagg	660
ggtcctgtg tgacttggtt ttccgtccaa ttccccaaagt aggac	706

<210> 64

<211> 630

<212> DNA

<213> Homo sapien

<400> 64

atgctgctcg agcggcgcag tggtgatgga tcgtggcgc ggccgaggta catcgacttc	60
actgcagacc aggtggacct gacttctgct ctgaccaaga aaatcactct taagacccca	120
ctggtttcct ctcccatgga cacagtacaca gaggctggga tggccatagc aatggcgctt	180
acaggcggta ttggcttcat ccaccacaaac tcctaagtat atgattgcga gtggaaaaat	240
aggggacaga aatcaggtat tggcagtttt tccatttca tttgtgtgtg aatttttaat	300
ataaaatgcgg agacgtaaag cattaatgca agttaaatg tttcagtgaa caagtttcag	360
cggttcaact ttataataat tataaataaa cctgttaaat ttttctggac aatgccagca	420
tttggatttt ttaaaacaa gtaaatttct tattgtggc aactaaatgg tgtttgtagc	480
attttatca tacagtagat tccatccatt cactatactt ttcttaactga gttgtcctac	540
atgcaagtac atgttttaa tggtgtctgt ctctgtgt gttcctgtaa gtttgttatt	600
aaaatacatt aaactataaa aaaaaaaaaa	630

<210> 65

<211> 4247

<212> DNA

<213> Homo sapien

<400> 65

gcccccttcta ggattttcta tgcacatgca catatatcta tgttttaaat cacagatggc	60
tgtttactgt ataacctgtc ctgcacccctg cttttcaact taatttgcata agaatcatcc	120
catatctgca caccattcct aaagcaacac aattttccat tatgatcatt agctagttcc	180
cactgacgag gacctgcgtc cttccgaac ctggccatca caaatgcggc tgcaagcagt	240
gtccttgcgtt ttaggtgtt ctgcataatgt gtaattgttag ccgtcagatt gctgacatag	300
atttgcgtgt cacagggtag tgcatactgtt tgaaatgtt gttggctgttg tcacactacc	360
ctttaaggtg atgtcaggct gtgccttgac tagccataca tgggagttct gttctcccac	420

accccaccaa cagtgggatc ttcataatga cccattatgg tccaaggagg cattgaatta	480
cagccactgc agagggctcc acatggctca cctggcccca ccaggccacg agtgtttga	540
ggcccttggg gccaatcttc caatccttga gctgcgttga caagccagtg cgtggctcg	600
gtgtgctgca gggttgagag agggctggtc atttagagtc agtatccaag actggattaa	660
tacaagaccg ttttggttt ttaggacaga ggaagagaag aaagaatgga ttcaggtgaa	720
gcttctaaag ttgtaaagta accaaatgac caagtgattg agcagtaagg ccattcatgt	780
tggttcagag ggtgggggccc ctggcctccg ggactgctgg ccctgtgctg taccctgcag	840
ggatgcagtg accacactgt gccttcataa gcagcttcag atgccacaag cccttcaacc	900
cattcatttc tttgagggcc cccaaataaa catgagcggg cctgggtggag tcacaggcca	960
ggttccccgc tcagggtgga tctcctcaat ctggacagct ccaaggggag accacatcat	1020
tggtaggggg gaagggagat ggccagtggc ctgggcattt tgctggaaac gccaaggccc	1080
gtcctcggga cagaggcagg cgcgccctc ctgtgggttag catcctccca tgcccttta	1140
tagtcctcac tgggtgttg ctgtgtccag atcatccagg ccaccatcga gaagcacaaa	1200
cagaacagcg aaaccccaa ggctttgggt ggcccttca gccaggatga ggaccccagc	1260
ctctctccag acatgcctat cacgagcacc agccctgtgg agcctgtggt gaccaccgaa	1320
ggcagttcgg gtgcagcagg gctcgagccc agaaaaactat cctctaagac cagacgtgac	1380
aaggagaagc agagctgtaa gagctgtggt gagaccccaa actccatcac caagaggagg	1440
catcaactgca agctgtgtgg ggccgtcatc tgtggaaagt gctccgagtt caaggccgag	1500
aacagccggc agagccgtgt ctgcagagat tgtttcttca cacagccagt ggccctttag	1560
agcacagagg tgggtgctcc cagtccttgc tccctcttgc gtggcgccgc agaccccttca	1620
gacacctgtt cctgtgtcccc agcagctcta gtcgccttgc ttccggagt gtccctggga	1680
ccaggataga tgtgggtgtg tctcagtggtt gccccccaggc tgggaacaca ctcggaaagat	1740
cccgtgtgtg cacgctggct tctcgggacc agccaggccc aaggcttggg aaccatttgc	1800
atccaggagt tagatggaa caagtgtcac tcctggtcac ctgggtggag ggcattggatt	1860
gaagctgccc tcaagtcacgc gccttgaggc ccacctgagg ggagatactg cccttcttca	1920
tgcctgcctc acctccagct gtcagagggt gaggcaggccc ctggaggcag gcacccccc	1980
ctcctcttcc ccttcccttcc tccctcttcc tccctcttcc ttctcttcc ttctcttcc	2040
ccttcttcc cctccgcctc gtcgttttc tccttcttcc tccctcttcc tccctgttcc	2100
ttctcttcc cgttccctt ttctacttcc ccctcttcc tccctcttcc tttctcttcc	2160
tcctcttcc cgttcccttcc tcctcttcc ttctcttcc tccctcttcc tccacttcc	2220

82

ctttctgaga	gtagctgctg	ctttccctga	cttccttcc	caagaataca	gtctattctg	2280
gggggttgca	ggacagctga	aaaagaaaac	gtgagttcc	agaaaagcac	atcgccctt	2340
ctctccactc	gacacagcct	gcgggttgca	gtgccagctg	cacgttggag	tctcctggga	2400
agcttctagt	gttcatctcc	ctccttagagg	ctcagtgagg	tgatcggggg	gtcagtgtgt	2460
tttcagagtc	ccagctgctc	tgacttacag	ctgtggttgg	agccctgacc	tagtccttg	2520
gtctaagaag	tcaaggatgt	aagtctatta	ggactgagtg	ttgaatggtc	ttgatggta	2580
agaaggacct	tgcactgcct	accgagctcc	ctcatgcctg	caatttaaat	accgaggtgt	2640
gttctgagag	ccccaggtcc	ttggggcacg	ggcctccagg	gccctgttaag	gtggactctg	2700
gcctcgaagg	ggggcccagg	tctgaggggt	ccagctctgg	cagggggtgt	ctaagccacg	2760
tccaaatggt	ggacagtagc	tccaaacagt	acatgacaga	gggagggcct	gtgtgtgtgg	2820
tgtgtgtgt	tgtgctgtgt	gtgcacgcac	gtgcatgtat	atggtgtgtg	tagaactgac	2880
acgcttgaag	tccctgtccc	gtgaccacc	ctcaggtact	tgctgttggg	ttttctatg	2940
attttgcgt	gtcctgaccc	agcatatgga	tgcctcccc	ttggagggcc	cagcctgcaa	3000
accccatctg	ataccacaca	tgcctatgtc	ctgcaggctt	ctagtgcaccc	ttgcaagccg	3060
ggtgaggctg	gtcccagtct	ctgtgcagat	cttggccatg	tctccacagc	cagctcccc	3120
caagcagtgg	ccctggccct	accccagct	cctcagaaca	ggggtgacca	ctgcagggga	3180
gactgcccct	aacctgtgtc	tttgcgtccc	cagaagacac	ccactgcaga	cccccagccc	3240
agcctgctct	gccccccct	gcggctgtca	gagagcggtg	agacctggag	cgaggtgtgg	3300
gccgccatcc	ccatgtcaga	tccccaggtg	ctgcacctgc	agggaggcag	ccaggacggc	3360
cggctgcccc	gcaccatccc	tctccccagc	tgcaaactga	gtgtgccgga	ccctgaggag	3420
aggctggact	cggggcatgt	gtggaagctg	cagtgggcca	agcagtcctg	gtacctgagc	3480
gcctcctccg	cagagctgca	gcagcagtgg	ctggaaaccc	taagcactgc	tgcccatggg	3540
gacacggccc	aggacagccc	ggggccctg	cagttcagg	tccctatggg	cgcagctgct	3600
ccgtgagctg	agtctcccac	tgccctgcac	accaccacat	tggacctgtg	ctgtcctggg	3660
aggtgggttt	ggaggccccca	tgaagagcgc	cctggactgc	tgagggtggg	ccaacagccc	3720
agagctcagg	acacttggct	ttgggggaa	ggaaactgag	gcccagagag	ggcaaccac	3780
tggccaagg	tcacccagca	agttttggct	aagagctgg	cctccagccc	cagcagtgtg	3840
gcccagagca	ggggccgact	gccaaggataa	ccatcatcca	tatggccgt	gtggtgatgc	3900
tggcccgaa	ggcagaaaga	ggcagcatgg	gcactgccag	ggacagccac	atcctgctgg	3960
tctgcagcgt	ggccaccccc	gcctctgccc	agcctgtcta	caccgtgtga	gctgaatcgt	4020
gacttgcctc	ccacccctt	tctctgtcct	ctcctgaggt	tctgcctgca	gcccccagga	4080

ggtgtggcctg ccccatccta gctggactca tggttcctaa ataaccacgc tcagaagctc	4140
tgcttaggact tacccccagcc actgagtgcc aggcgcatga gatttgtggc tgttcctgat	4200
gctagtgcc cacagtgcctt atctgcataa ataaacactg gccacca	4247
<210> 66	
<211> 513	
<212> DNA	
<213> Homo sapien	
<400> 66	
ctcttagagga tctcggtgc ctgcacttc mmtctgacts tatacgctggc cattcctacc	60
tcggaggtgg aggccggaaa ggtcgcacca agagagaagc tgctgccaac accaaccgccc	120
ccagccctgg cgggcacgag aggaaactgg tgaccaagct gcagaattca gagaggaaga	180
agcgaggggc acggcgctga gacagagctg gagatgaggc cagaccatgg acactacacc	240
cagcaataga gacgggactg cggaggaagg aggacccagg acaggatcca ggccggcttg	300
ccacacccccc cacccttagg acttattccc gctgactgag tctctgaggg gctaccagga	360
aagcgcctcc aaccctagca aaagtgcag atggggagtg agaggctggg aatggagggg	420
cagagccagg aagatccccc agaaaagaaa gctacagaag aaactggggc tccctccaggg	480
tggcagcaac aataaataga cacgcacggc agc	513
<210> 67	
<211> 1772	
<212> DNA	
<213> Homo sapien	
<400> 67	
tgggcgtggac tcagggaccg actcttcccg tctcatgact gtgtttactg ggctggattt	60
tgggaagggg ccagattgca tcagacaggg cctgatggc tggagccaga ctgtggctcg	120
aggaggagac acagccttat aagctgaggg agtggagagg cccggggcca ggaaagcaga	180
gacagacaaa gcgttaggag aagaagagag gcagggaaaga caagccaggc acgatggcca	240
ccttccacc agcaaccaggc gccccccaggc agccccccagg cccggaggac gaggactcca	300
gcctggatga atctgaccc tatagcctgg cccattccta cctcggtaag gcccactcag	360
ccatetccac ggtccttcct cctctcccg aatcaggacc caccctctt gtttctctc	420
atttcctttc ctttcctctt cgtttcttc ttctttttt ttgagagag tcactctgtc	480
acccaggctg gagtgcagta gtgtgatcac aacaactcaa acaactcacc gcagcctcga	540
actcctgggc tcaagtgatc ctccctgcctt agcctcctga gtggctggaa cttcagggtgt	600
acaccacactg cagtggtgag atggggtctc actatgtttc ccaggctgat cccaaattcc	660

tgggctcaag caatttcctt gccttggcct cccaaagtgc tgagattaca ggagttagcc	720
accctgtccca gcccactcac cttttctag ccaacctgtt cttggaccc tcacgtcacc	780
cctgtctaat cccttatccc aggagtgtca tttactcag cttggacct cacacacatc	840
tgggtccca cattccacag aggggaagca gcaggcttcc ctgtcttccc cccatcccc	900
caaccctgaa cccctgcctc tcctctgaca gggctctca tcatgcctat gcccacttca	960
cctctgactc ctgccttggt tacaggaggt ggaggccgga aaggtcgac caagagagaa	1020
gctgctgccca acaccaaccg ccccaaccct ggccggcacf agaggaaact ggtgaccaag	1080
ctgcagaatt cagagaggaa gaagcgaggg gcacggcgct gagacagagc tggagggtaa	1140
ggagtccgggg ggcccagaga gctcaaggtg gtgttctgc catgaaggac aggccggaag	1200
gtgtgtgatt gggtggggag gagggatcag gcaactgttgc ttgtatgca gaataaaacg	1260
agacatatgt ttgattgtga gtttccttagt ggccagagca aagtgggaac acagaacctt	1320
tccaattgaa gggaaatttgc acttacagag acagaaatttgc aggatagaga gggtgggtcc	1380
ttccctggag tcacaaatca agtgagtggg agaggcagaa tttagaacccca gatctctgtc	1440
ccttttacca cctgttttc ctcacccccc agatgaggcc agaccatgga cactacaccc	1500
agcaatagag acgggactgc ggaggaagga ggacccagga caggatccag gcccgttgc	1560
cacacccccc acccttagga cttattcccg ctgactgagt ctctgagggg ctaccaggaa	1620
agcgccctcca accctagcaa aagtgcaga tggggagtga gaggctggga atggaggggc	1680
agagccagga agatccccca gaaaagaaag ctacagaaga aactggggct cctccagggt	1740
ggcagcaaca ataaatagac acgcacggca gc	1772

<210> 68
<211> 1864
<212> DNA
<213> Homo sapien

<400> 68	
tagatgcattt tcgagcggcg cagtgtatg gatgtggcg cggccgagggt ggggagctga	60
atcccgaaat atccccacat cgatgaaagc aaagcgaagc caccaagccca tcattatgtc	120
cacgtcgcta cgagtcagcc catccatcca tggctaccac ttgcacacag cctctcgtaa	180
gaaagccgtg ggcaacatct ttgaaaacac agaccaagaa tcactagaaa ggctttcag	240
aaactctgga gacaagaaag cagaggagag agccaaagatc atttttgcac tagatcaaga	300
tgtggaggag aaaacgcgtg ccctgtatggc cttgaaagaaag aggacaaaag acaagcttt	360
ccagtttctg aaactgcggc aatattccat caaagttcac tgaagagaag aggatggata	420
aggacgttat ccaagaatgg acattcaaag accaagttag tttgtgagat tctaacagat	480

85

gcagcatttt	gctgctaccc	tacaagcttc	tcttctgtca	ggactccaga	ggctggaaag	540
ggaccgggac	tggaaaggga	ccaggactga	acagactggt	tacaaagact	ccaaacaatt	600
tcatgccctg	tgctgttaca	gaggagaaca	aatgcttc	agcaaggatt	tgaaaactct	660
tccgtccctg	cagggaaaggga	ttgatgctga	tagaagagcc	tggacagatg	taatgagaac	720
taaagaaaac	agatggctgg	agatgacatt	tatccagggt	cactttgtca	ggcccttagga	780
cttaaatcga	agttgaacct	ttttttttt	ttaacccaaat	agataggggga	agggaggagg	840
gagagggagg	acagggagag	aaaataccat	gcataaattt	tttactgaat	ttttatatct	900
gagtgttcaa	aatatttcca	agcctgagta	ttgtctattt	gtatagattt	ttagaaatca	960
ataattgatt	atttatttgc	acttattaca	atgcctgaaa	aagtgcacca	catggatgtt	1020
aagttagaaat	tcaagaaagt	aagatgtctt	cagcaactca	gtaaaacctt	acgcccacctt	1080
ttgggtttgta	aaagggtttt	tatacatttc	aaacagggtt	cacaaaagtt	aaaataatgg	1140
ggtcttttat	aaatccaaag	tactgtaaaa	acattttaca	tatTTTTAA	atcttctgac	1200
taatgctaaa	acgtaatcta	attaaatttc	atacagttac	tgcagtaagc	attaggaagt	1260
gaatatgata	tacaaaatag	tttataaaga	ctctatagtt	tctataattt	atTTTactgg	1320
caaatgtcat	gcaacaataa	taaattattt	taaactttgt	ggcttttgtt	ctgtgtatgt	1380
tggctctaaa	ggaaaaaaaata	agatggtaaa	tgttgatatt	tacaaacttt	tctaaagatg	1440
tgtctctaac	aataaaagtt	aatttttagag	tagttttata	ttaattacca	aacttttca	1500
aaacaaattc	ttacgtcaaa	tatctggaa	gtttctctgt	cccaatctta	aaatataaaaa	1560
tatagatata	gaagttcata	gattgactcc	ttggcatttc	tatTTATGTA	tccattaagg	1620
atgagttta	aaaggctttc	tcttcatact	tttgaaaaat	ttcttctatg	attacagtag	1680
ctatgtacat	gtgtacatct	atTTTCCCA	agcaatatgt	tttgggttta	gagtctgagt	1740
gatgaccaag	attctgtgtg	ttactactgt	ttgtttaata	ggaacaaata	tagaaataat	1800
attatctctt	tgcttatttc	ccgttaaaac	tataataaaa	tgtttctaaag	acagcatacg	1860
taaa						1864

<210> 69
<211> 1572
<212> DNA
<213> Homo sapien

<400> 69	agatgctgt	cgagcggcgc	agtgtgtatgg	atgggcaggt	aaagggagct	gaattccgga	60
	agatccccac	atcgatgaaa	gcaaagcgaa	gccaccaagc	catcatcatg	tccacgtcgc	120
	tacgagtcag	cccatccatc	catggctacc	acttcgacac	agcctctgt	aagaaagccg	180

86

tgggcaacat	cttgaaaac	acagaccaag	aatcaactaga	aaggctttc	agaaactctg	240
gagacaagaa	agcagaggag	agagccaaga	tcattttgc	catagatcaa	gatgtggagg	300
agaaaacgcg	tgccctgatg	gccttgaaga	agaggacaaa	atgcttcag	caaggatttg	360
aaaactcttc	cgtccctgca	ggaaaggatt	gatgctgata	gaagagcctg	gacagatgta	420
atgagaacta	aagaaaacag	atggctggag	atgacattt	tccagggtca	ctttgtcagg	480
cccttaggact	taaatcgaag	ttgaactttt	ttttttttt	aaccaaata	ataggggaag	540
ggaggaggga	gagggaggac	agggagagaa	aataccatgc	ataaaattgtt	tactgaattt	600
ttatatctga	gtgttcaaaa	tatttccaag	cctgagtatt	gtctatttgt	atagattttt	660
agaaatcaat	aattgattat	ttatttgcac	ttattacaat	gcctgaaaaa	gtgcaccaca	720
tggatgttaa	gtagaaattc	aagaaagtaa	gatgtcttca	gcaactcagt	aaaaccttac	780
gccacctttt	ggttgtaaa	aggttttta	tacatttcaa	acaggttgca	caaaagttaa	840
aataatgggg	tctttataa	atccaaagta	ctgtaaaaac	atttacata	tttttaaat	900
cttctgacta	atgctaaaac	gtaatcta	taaatttcat	acagttactg	cagtaagcat	960
taggaagtga	atatgatata	caaaatagtt	tataaagact	ctatagttc	tataatttat	1020
tttactggca	aatgtcatgc	aacaataata	aattattgtt	aactttgtgg	ctttggtct	1080
gtgatgcttgc	gtctcaaagg	aaaaaataag	atggtaaatg	ttgatatttta	caaactttc	1140
taaagatgttgc	tctctaacaa	taaaagttaa	ttttagagta	gttttatatt	aattacaaa	1200
cttttcaaa	acaaattctt	acgtcaaata	tctggaaagt	ttctctgtcc	caatcttaaa	1260
atataaaata	tagatata	agttcataga	ttgactcctt	ggcatttcta	tttatgtatc	1320
cattaaggat	gagttttaaa	aggctttctc	ttcatacttt	tgaaaaattt	cttctatgat	1380
tacagtagct	atgtacatgt	gtacatctat	tttcccaag	caatatgttt	tgggtttaga	1440
gtctgagtga	tgaccaagat	tctgtgtgtt	actactgttt	gtttaatagg	aacaaatata	1500
gaaataatata	tatcttttgc	tttatttccc	gtaaaaacta	taataaaatg	tttctaagac	1560
agcatacgtaa	aa					1572

<210> 70
<211> 1265
<212> DNA
<213> Homo sapien

<400> 70
agatgcatgt cgagcgggcc gcagtgttga tggatacaag gccgtgaggt tctccagccc 60
ctccagagca ttgttgggaa cccgtgagat ctgggtgtgg tcgagatgga gctctctcac 120
accccacaga gctaatactaa atcttgcgt agaaaaagca ttctctaact ctacccacc 180

ctacaaaatg catatggagg taggctgaaa agaatgtaat ttttatttc taaaaatacag	240
atttgagcta tcagaccaac aaacccccc cctgaaaagt gagcagcaac gtaaaaacgt	300
atgtgaagcc tctcttgaat ttcttagttag caatcttaag gctcttaag gtttctcca	360
atattaaaaa atatcaccaa agaagtcctg ctatgtaaa aacaaacaac aaaaaacaaa	420
caacaaaaaa aaattaaaaa aaaaaacaga aatagagctc taagttatgt gaaatttgat	480
ttgagaaaact cggcatttcc tttttaaaaa agcctgttc taactatgaa tatgagaact	540
tctaggaaac atccaggagg tatcatataa cttttagaa cttaaatact tgaatattca	600
aatttaaaag acactgtatc ccctaaaata tttctgtatgg tgcaactactc tgaggcctgt	660
atggccccctt tcatcaatat ctattcaaata atacaggtgc atatatactt gttaaagctc	720
ttatataaaa aagccccaaa atattgaagt tcatctgaaa tgcaagggtgc tttcatcaat	780
gaaccttttc aaacttttct atgattgcag agaagcttt tatataccca gcataacttg	840
gaaacaggta tctgacctat tcttatttag ttaacacaag tgtgattaat ttgatttctt	900
taattccta ttgaatctta tgtgatatga tttctggat ttacagaaca ttagcacatg	960
taccttgtc ctccccattca agtgaagttta taatttacac tgagggttc aaaattcgac	1020
tagaagtgga gatatattat ttatattgc actgtactgt atttttatat tgctgtttaa	1080
aacttttaag ctgtgcctca cttattaaag cacaaaatgt tttacctact ccttatttac	1140
gacgcaataa aataacatca atagatttt aggctgaatt aatttgaag cagcaatttg	1200
ctgttctcaa ccatttttc aaggctttc attskwcaaa kwwaataaaaw martagayww	1260
twarg	1265

<210> 71
 <211> 7232
 <212> DNA
 <213> Homo sapien

<400> 71	
tcctccctcc ctccccatcc ccaatcctga ttcctgttta caaagaatgt taaaaacaaa	60
ggaattatgt ataacagttc ccagttgtc cagggaaattc tcagattata aagagacatt	120
acaaaatgaac aagtgaagag aagaaccttg gtggccaa catagtatgg ccatttttt	180
atactcaaaa tatagaaaga caacctcaga ataagaaaac ttttggatg gaataaatca	240
agtttatcat taaaatgca agaaaaaaaaac tctccaaatg ttgctgatct tctgttttaa	300
actactgtta gaccggagaa gcggagagca gggaaatccg ccaaagagtt ttggatgaaa	360
attaatcagc cctgtctacc gtagtcacac cccactgccc ttgagaccca atccttcgga	420
aggagtgtcc aagaggtata aagcaaaaacg aaaaaacagt tcgcaaattc cagagttcgt	480

88

tttctctcat	taaaaatata	aatatcaggc	taacacatgt	tgacacaccaa	taacaggac	540
acagaatccc	tcctggaaga	ccgacgggcc	cacggacccc	acgggtgcc	cggtggtgga	600
cgaggtaag	taacttggtt	cagggtgtct	gggcacacac	ctgcgtgaga	ctctgtctct	660
gctgctcctc	tcatctctac	gccgattct	ccccacaatc	ctccctttc	cttgggcccc	720
cgacgcctct	ccgaccaaca	gtctccccag	ccccgcagct	tctcttttc	agacctttac	780
ttcttgatcc	tcactccata	gtgagatgtg	gccttcage	aaataaattg	tgctcaggga	840
gacagccaat	tgtcccttgc	cgtcctcctg	agggtgcctg	gagcttaagc	actgtgtgct	900
cttggcctcc	acactgggga	tgccgctgac	tcccactgtc	cagggcttcc	agtggattct	960
ccgaggccct	gatgttagaaa	cttcccatt	gggtgcacca	agagcagcct	cacatggtgt	1020
gggctgacat	caagagctgc	cagatccaac	aggaagatgg	ccaatcttc	ctaagctgct	1080
cacccatcaa	gaaaacgaat	cgtactgcta	agaattcaaa	cttcagcagt	catggggagc	1140
cttggaaagga	gcccgaaatca	ctgatggaat	tggacagtgc	atggagatgg	ttcagcagga	1200
caagggttaag	tgcaggggca	agtccaggc	atactgagag	acaacgagt	gchgctgacag	1260
agacagacaa	agataaaatc	aaaagttgt	gcttcatttt	caaaaactca	aactaataac	1320
aaacttggcc	ttatgagaaa	taataagtat	ttttctat	acatgagaat	ttaatctcaa	1380
aacaggaatc	agaaacat	taagtccagg	gcataaaacc	taaaccactg	ctcatat	1440
ttctttctaa	atagagcaaa	gtgtaaaatc	ttctccat	aatgcacatt	gtgcttat	1500
aaaggccagt	cttagtgaga	atcattggta	ttccatagaa	gagtgaatta	aacacagcca	1560
agggaaagacc	caagtctcat	acttctcttg	tatattccag	agttccagg	gaattccagg	1620
tgatagaggt	gatctccat	actgttaaag	caagggtgca	gacacttgg	aattttggc	1680
ccagtaact	aggaggcac	acctctgtcc	tggaaaatac	tacaggaat	tatactctc	1740
ctatgactca	ttctggtcat	tottccagca	tcacaaaaac	caaaaaaaaaa	aaaggaaata	1800
tgtccaaata	catgatttgc	tatccctcct	cttcagg	tttacctgtt	acttacggat	1860
aacagcatta	ccacaggatt	atgatgaaga	tacaatgtcc	aaatataa	acagtttga	1920
gcaaaatgcc	ttgtacgaat	ttgtcaatga	acaactagta	aataattat	tgaatattt	1980
ctgaattata	tggatcctat	gaataattac	tgaataattc	atgtgattgc	ttttattggc	2040
agtgctgaaa	actcateccc	gtgtgac	aagtaagcca	tgtactctg	tgaacctgca	2100
gttttatcat	ttttaaaaata	aagaaacatg	acagattt	tttgcacac	agaatgtcag	2160
gtctccaga	tgccagaaaa	tacatttact	taaagccgtt	gatacgtt	aaagccgtt	2220
ccttacagt	tcattggagg	acagtgtgga	gtgcagagag	acatgctt	aaatgggatt	2280
gatccagtcc	tccttccttc	actaccacat	gaatgctggg	cagcccaggg	tcaacccacc	2340

gcaccctcaa	ctcaggcaag	tccagcagcc	aatcttagga	gacctggct	acagaacagt	2400
ctccccaa	tttccatgtt	tttttttttttt	tttttttttttt	tttttttttttt	tttttttttttt	2460
gttcagggaa	tcactctttc	ctactcattc	ctctcatctc	aaactcacct	tctactgcaa	2520
cactgaggat	caccaaccaa	ccgtgaccat	aaccttgatc	ttgccatgtt	ctgttagtgg	2580
aatgcaaccc	aaaatcaatg	gtgttagtc	atctgaacaa	aatatatatc	aaaccatatt	2640
gcataagaac	cgctcatggc	cctgttcttt	tcagtatatg	ggaaaacaaa	atggaaacaa	2700
caaaatagca	tcaggtttat	gaaacttccc	aagatagatg	gtcacacatg	ttttcaggag	2760
atctctatat	aaatgatttt	gatcaactga	taccttgaaa	agagctctg	tgacactaga	2820
atgacatcca	taagtgacaa	gtataaaatg	tagcgctcag	tgacatcaaa	aaccaaatca	2880
acccacatag	aggaagagct	ctggacatag	ggatgtcaaa	ctggtctaga	gtgtaatgaa	2940
aagcaaagat	ggtgccccag	tgagaaaaaa	gaaatcaaca	taacaatggg	aaacagcaag	3000
aagaatactg	agacagggaaa	gacaacattt	tttacaaatg	aattattcat	tcactttcta	3060
gtggatacag	acaaaactgc	agaagaccca	gaggaaatca	gggcaggcta	aaagtttgat	3120
atcttacacc	tgtggaaaag	gccttaagct	ctgttttaac	tgagagcagg	tgggttgact	3180
tcatgactac	cattaagaaa	atacaacctg	ttggaaaact	gtttctgcct	tgtatgtgtt	3240
gtacagacaa	gagataaaca	gtgaggaata	tgcttagatg	tattggaaa	gacacgggtc	3300
tgtggcatca	tcacaagggt	acacgaatac	tgagagtgaa	tgctgaagga	atgatcccc	3360
ttgggtgtga	ccctcaggtg	agactagggt	gcctgtgttt	caggaaagcc	tggcaattg	3420
gaatgcaggg	ctcctaagat	tccatgacac	ccccaccttc	taattctgtt	attgcaactg	3480
cagacggta	cctggcacgc	tggccacagt	ctacctca	cttatcagag	tctgagctac	3540
tggcagtgt	ttcagctctg	agttcaggca	cctcgAACCT	tgtttttgtg	gtgaaggatc	3600
ctaaagtgt	gtggggagtg	atcacattt	tcacaacatc	cctggctcca	cctcttctgc	3660
cacaaacgtc	agcatggtg	tatcagctgg	cccttggtcc	agcgagaagg	cagagacgaa	3720
cattttagaa	atcaacgaga	aattgcgccc	ccagctggca	gagaacaaac	agcagttcag	3780
aaacotcaa	gagaaatgtt	ttgtaactca	actggccggc	ttcctggcca	accgacagaa	3840
gaaatacaag	tatgaagagt	gtaaagacct	cataaaattt	atgctgagga	atgagcgaca	3900
gttcaaggag	gagaagctt	cagagcagct	caagcaagct	gaggagctca	ggcaatataa	3960
agtccctggtt	cactctcagg	aacgagagct	gaccctgtt	agggagaagt	tacggaaagg	4020
gagagatgcc	tcccgctcat	tgaatcagca	tctccaggcc	ctcctca	cgatgagcc	4080
agacaagtcc	caggggcagg	acctccaaga	acagctggct	gaggggtgt	gactggcaca	4140

90

gcacccttgtc	caaaagctca	gcccgaaaaa	tgacaacgt	gacgatgaag	atgttcaagt	4200
tgagggtggct	gagaaagtgc	agaaatcgtc	tgcccccagg	gagatgccga	aggctgaaga	4260
aaaggaagtc	cctgaggact	cactggagga	atgtgccatc	acttgttcaa	atagccatgg	4320
cccttatgac	tccaaccagc	cacataggaa	aacccaaaatc	acatttgagg	aagacaaaagt	4380
cgactcaact	ctcattggct	catcctctca	tgttgaatgg	gaggatgctg	tacacattat	4440
cccagaaaaat	gaaagtgtat	atgaggaaga	ggaagaaaaa	gggccagtgt	ctcccaggaa	4500
tctgcaggag	tctgaagagg	aggaagtccc	ccaggagtcc	tggatgaag	gttattcgac	4560
tctctcaatt	cctcctgaaa	tgttgcctc	gtaccagtct	tacagccgca	catttcactc	4620
attagaggaa	cagcaagtct	gcatggctgt	tgacataggc	ggacatcggt	ggatcaagt	4680
gaaaaaggag	gaccaagagg	caacaggtcc	cagccaggct	cagcaggag	ctgctggatg	4740
agaaagggcc	tgaagtcttg	caggactcac	tggatagatg	ttattcaact	ccttcagggt	4800
atcttgaact	gactgactca	tgccagccct	acagaagtgc	ctttacata	ttggagcaac	4860
agcgtgttgg	ctgggctctt	gacatggatg	aaattgaaaa	gtaccaagaa	gtggagaag	4920
accaagaccc	atcatgcccc	aggctcagca	ggagctgct	ggatgagaaa	gagcctgaag	4980
tcttgacgga	ctcaactggat	agatgttatt	cgactccttc	aggttatctt	gaactgcctg	5040
acttaggcca	gccctacaga	agtgctgttc	actcattgg	ggaacagtagc	cttggcttgg	5100
ctcttgacgt	ggacagaatt	aaaaaggacc	aggaagagga	agaagaccaa	ggccaccat	5160
gccccaggct	cagcaggag	ctgctggagg	cagtagagcc	tgaagtcttg	caggactcac	5220
tggatagatg	ttattcaact	ccttcagtt	gtcttgaaca	gcctgactcc	tgccctgcct	5280
atggaagttc	cttttatgca	ttggagggaaa	aacatgttgg	ctttcttctt	gacgtggag	5340
aaattgaaaa	gaaggggaag	gggaagaaaa	gaaggggaag	aagatcaacg	aagaaaagaa	5400
ggagaagggg	aagaaaagaa	gggaaagaag	atcaaaaccc	accatgcccc	aggctcagca	5460
gggagctgct	ggatgagaaa	gggcctgaag	tcttgacgga	ctcaactggat	agatgttatt	5520
caactccttc	aggttatctt	gaactgactg	actcatgcca	gccctacaga	agtgcgttt	5580
actyattkga	gsaacagcry	rtsagcttc	gcccttgacg	tggacaatag	agtttcttta	5640
ctttgatggg	aakaaggct	ccacctgagt	cttccagatg	ggagtcatat	tcccacagta	5700
agcagccctt	actaagccga	gagatgtcat	tcctgcaggc	aggacctata	ggcacgtgaa	5760
gatttgaatg	aaactmtagt	tccayttgga	agcccagrca	wrggatgggt	cagtgrgcak	5820
ggctctmttc	ctaktctcag	rccatgccwg	tggcamcctg	tgctcagtct	gaagacaatg	5880
gacccaagtt	aggtgtgaca	cgttcacata	actgtgcagc	acatgccggg	agtgatcagt	5940
cagacatttt	aatttgaacc	acgtatctct	ggtagctac	aaagttcctc	agggatttca	6000

ttttgcaggc atgtctctga gcttctatac ctgctcaagg tcagtgtcat ctttgtttt 6060
 agctcatcca aagggtttac cctgggttca atgaacctaa cctcatttt tgtatcttca 6120
 gtgttgaatt gtttagctg atccatctt aacgcaggag ggatccttg ctgaggattg 6180
 tatttcagaa ccaccaactg ctcttgacaa ttgttaaccc gctaggctcc tttggttaga 6240
 gaagccacag tccttcagcc tccaattggt gtcaagtactt aggaagacca cagctagatg 6300
 gacaaacagc attggggagac cttagccctg ctccctctcgat ttccatcttg tagagaacag 6360
 gagtcaggag cccgctggcag gagacagcat gtcacccagg actctgcgg tgcaaatat 6420
 gaacaacgcc atgttcttgc agaaaacgct tagcctgagt ttcataggag gtaatcacca 6480
 gacaactgca gaatgttagaa cactgagcag gacaactgac ctgtctcctt cacatagtcc 6540
 atatcaccac aaatcacaca acaaaaagga gaagagatat tttgggttca aaaaaagtaa 6600
 aaagataata tagctgcatt tcttttagtta tttgttaaccc caaatatttc ctcatcttt 6660
 tgttgttgc attgatggtg gtgacatgga cttgtttata gaggacaggt cagctgtctg 6720
 gctcaagtat ctacattctg aagttgtctg aaaatgtctt catgattaaa ttcaagcctaa 6780
 acgttttgc gggAACACTG cagagacaat gctgtgagtt tccaaacctt gcccatctgc 6840
 gggcagagaa ggtcttagttt gtccatcagc attatcatga tatcaggact ggttacttgg 6900
 ttaaggaggg gtcttaggaga tctgtccctt ttagagacac cttacttata atgaagtatt 6960
 tgggggggtg gttttcaaaa gtagaaatgt cctgtattcc gatgatcatc ctgtaaacat 7020
 tttatcattt attaatcatc cctgcctgtg tctatttata tattcatatc tctacgctgg 7080
 aaactttctg cctcaatgtt tactgtgcct ttgttttgc tagttgtgt tggtaaaaaa 7140
 aaaaacattt tctgcctgag ttttaatttt tgtccaaagt tattttatc tatacaatta 7200
 aaagcttttgc cctcttagatc gcggggcgcc gc 7232

<210> 72
 <211> 6876
 <212> DNA
 <213> Homo sapien

<400> 72
 cggggcctgt gtcccccgcg ctggattttt cgccgcgc tgccgcgc agcccaactc 60
 tcgtggcgcc tggggaaagaa actcgctggc ggggtttctg tggcatccca ggggggtggag 120
 ggacggagca gcttcggggg cacgtccccc tatatcctgt agaggacact gacccgcac 180
 cccaccctcc agggcagaaa tccgtccctt ctgcggaccc gagaggcgag cgcgctcgcg 240
 cccctgactt gcaaagttgg ggtctttact ggccctccggg cttctgctcc tggcggttgc 300
 tccaggctgg tgatggcaa gccaggtgtg ccagctccag gatgcacatg agcagcattt 360

gtagccatcg ctgaatcacc tcctgactag cggggcaagc ctc当地atgaa ccgcaggatt	420
tccggcaatc tgaaggcaaa tcctgttag acccaggcga aggttcccgg tgacccggc	480
tctcaccagc caattgtccc ttgccgtct cctgagggtg cctggagctt aagcaactgtg	540
tgctcttggc ctcccacactg gggatgccgc tgactcccac tgtccagggc ttccagtgg	600
ttctccgagg ccctgtatgt aaaaacttccc catgggtgc accaagagca gcctcacatg	660
gtgtggctg acatcaagag ctgccagatc caacaggaag atggccaatc tttcctaagc	720
tgctcacctt acaagaaaaac gaatcgtaact gctaagaatt caaactttag cagtcatgg	780
gagccttgg aaggagccga atcaactgtg gaattggaca gtgcattggag atggttcagc	840
aggacaaggg taagtgcagg ggcaagtcca ggtcataactg agagacaacg agtggcgctg	900
acagagacag acaaagataa aatcaaaaatg ttgtgttca tcttcaaaaaa ctcaaaactaa	960
taacaaaactt ggccattatga gaaataataa gtatttttctt atttacatga gaatttaatc	1020
tcaaaaacagg aatcagaaac atattaagtc cagggcataa aacctaaacc actgctcata	1080
tttattcttt ct当地atagag caaagtgtaa aatcttctcc ataaaaatgca catttgctt	1140
atgaaaaggc cagtcattgt gagaatcatt ggtattccat agaagagtga attaaacaca	1200
gccaaaggaa gacccaaatgc tcataacttctt cttgttatatt ccagagtcc agggaaattc	1260
caggtgatag aggtgatctc ccatactgtt aaagcaaggt tgcagacact tggaaatttt	1320
ggtcccagta ctctaggagg tcacacctt gtcctggaaa atactacagg aatgtataact	1380
cttcctatga ctcattctgg tcattctcc agcatcacaa aaaccaaaaaa aaaaaaaaaagga	1440
aatatgtcca aatacatgtat ttgctatccc tcctttcag gtttcttacc tgttacttac	1500
ggataaacgc attaccacag gattatgtat aagatacaat gtccaaatataa aacacagtt	1560
ttgagcaaaa tgccttgtac gaattggtca atgaacaact agtaaataat tatgtgaata	1620
tttactgaat tatatggatc ctatgaataa ttactgaata attcatgtga ttgttttat	1680
tggcagtgtct gaaaactcat ccccggtgtga cctcaagtaa gccatgtAAC tctgtgaacc	1740
tgcagttta tcattttaa aataaagaaa catgacagat tttcattatg acacagaatg	1800
tcaggtctcc cagatgccag aaaatacatt tacttaaagc cggtgatacg tcttaaagcg	1860
gtttccttac agtgcattt gaggacagtg tggagtgcag agagacatgc tttgaaatgg	1920
gattgatcca gtcctcttc cttcaactacc acatgaatgc tggcagccc agggtcaacc	1980
cacccgcaccc tcaactcagg caagtccagc agccaatctt aggagacctg ggctacagaa	2040
cagtctccca agtccagggc tcacaaaacc taggtgggt gaaagctgag aaagcgagga	2100
gttggttcag gggatcactc tttcctactc attcctctca tctcaaaactc accttctact	2160

93

gcaacactga ggatcaccaa ccaaccgtga ccataacctt gatcttgcca tgttctgtta	2220
gtggaatgca acccaaaaatc aatggtgtt aatggatctga acaaaaatata tatcaaacca	2280
tattgcataa gaaccgctca tggccctgtt ctttcagta tatggaaaa caaaatggaa	2340
acaacaaaat agcatcagg ttatgaaact tcccaagata gatggtcaca catgtttca	2400
ggagatctct atataaatga ttttgcac ttgatacctt gaaaagagct cttgtgacac	2460
tagaatgaca tccataagtg acaagtataa aatgttagcgc tcagtgacat caaaaaccaa	2520
atcaacccac atagaggaag agctctggac atagggatgt caaactggc tagagtgtaa	2580
tgaaaagcaa agatggtgcc ccagtgagaa aaaagaaaatc aacataacaa tggaaacag	2640
caagaagaat actgagacag gaaagacaaac atttttaca aatgaattat tcattcactt	2700
tctagtgat acagacaaaaa ctgcagaaga cccagaggaa atcagggcag gctaaaagtt	2760
tgatatctta cacctgtgga aaaggcctta agctctgtt taactgagag caggtggggt	2820
gacttcatga ctaccattaa gaaaatacaa cctgtgggaa aactgtttct gccttgatga	2880
tgttgacag acaagagata aacagtgagg aatatgttta gatgtattgg gaaagacacg	2940
ggtctgtggc atcatcacaa gggtacacga atactgagag tgaatgctga aggaatgatc	3000
cccattggtg gtgaccctca ggtgagacta ggggcctgt gtttcaggaa agcctggca	3060
attggaatgc agggctccta agattccatg acacccccac cttctaattc tgttattgca	3120
actgcagacg gttacctggc acgctggca cagtctacct cactcttac agagtctgag	3180
ctactggcag tgcttcagc tctgagttca ggcacctcg aacctgtttt tgtggtaag	3240
gatcctaaag tgctgtgggg agtgatcaca ttttcaccaa catccctggc tccaccttt	3300
ctgccacaaa cgtcagcatg gtggatcatcg ctggcccttg gtccagcgg aaggcagaga	3360
cgaacatccc agaaatcaac gagaaattgc gccccagct ggcagagaac aaacagcagt	3420
tcagaaacct caaagagaaa tgtttgtaa ctcaactggc cggcttcctg gccaaccgac	3480
agaagaaata caagtatgaa gagtgtaaag acctcataaa atttatgtg aggaatgagc	3540
gacagttcaa ggaggagaag cttgcagagc agctcaagca agctgaggag ctcaggcaat	3600
ataaaagtccct ggttcactct caggaacgag agctgaccca gttaaaggag aagttacggg	3660
aagggagaga tgcccccgc tcattgaatc agcatctcca ggccttcctc actccggatg	3720
agccagacaa gtcccagggg caggacctcc aagaacagct ggctgagggg tgttagactgg	3780
cacagcacct tgcctccgc tcattgaatc agcatctcca ggccttcctc actccggatg	3840
aagttgaggt ggctgagaaa gtgcagaaat cgtctgcccc cagggagatg ccgaaggctg	3900
aagaaaagga agtcccgtag gactcactgg aggaatgtgc catcaactgt tcaaataagcc	3960
atggccctta tgactccaac cagccacata ggaaaaccaa aatcacattt gaggaagaca	4020

aagtgcactc aactctcatt ggctcatcct ctcatgttga atgggaggat gctgtacaca	4080
ttatcccaga aaatgaaagt gatgatgagg aagaggaaga aaaagggcca gtgtctccca	4140
ggaatctgca ggagtctgaa gaggaggaag tccccagga gtcctggat gaaggttatt	4200
cgactctctc aattcctcct gaaatgttgg cctcgatcca gtcttacagc ggcacatttc	4260
actcattaga ggaacagcaa gtctgcatgg ctgttgacat aggccggacat cggtggatc	4320
aagtaaaaaa ggaggaccaa gaggcaacag gtcccagcca ggctcagcag ggagctgctg	4380
gatgagaaaag ggctgaaagt cttgcaggac tcactggata gatgttattc aactccttca	4440
gttatcttg aactgactga ctcatgccag ccctacagaa gtgcctttta catattggag	4500
caacagcgtg ttggctgggc tcttgacatg gatgaaattt gaaatcacca agaagtggaa	4560
gaagaccaag acccatcatg ccccaggctc agcagggagc tgctggatga gaaagagcct	4620
gaagtcttgc aggactcact ggatagatgt tattcactc cttcaggta tcttgaactg	4680
cctgacttag gccagcccta cagaagtgt gttcaactcat tggaggaaca gtaccttggc	4740
ttggcttttgc acgtggacag aattaaaaag gaccaggaag aggaagaaga ccaaggccca	4800
ccatgccccca ggctcagcag ggagctgctg gaggcagtag agcctgaagt cttgcaggac	4860
tcactggata gatgttattc aactccttcc agttgtcttgc aacagcctga ctcctgcctg	4920
ccctatggaa gttcctttta tgcattggag gaaaacatg ttggcttttc tcttgcgtg	4980
ggagaaattt gaaaagaaggg gaaggggaag aaaagaaggg gaagaagatc aacgaagaaa	5040
agaaggagaa ggggaagaaa agaaggggaa gaagatcaa acccaccatg ccccaggctc	5100
agcagggagc tgctggatga gaaagggcct gaagtcttgc aggactcact ggatagatgt	5160
tattcaactc cttcaggta tcttgaactg actgactcat gccagcccta cagaagtgcg	5220
ttttactyat tkagsaaca gcryrttsag cttgcctt gacgtggaca atagatttc	5280
tttactttga tggaaakaag gtctccacct gagtcttcca gatgggagtc atattccac	5340
agtaagcagc ccttactaag ccgagagatg tcattcctgc aggcaggacc tataggcacg	5400
tgaagatttgc aatgaaactm tagttccayt tggaaagccca grcawrggat gggtcagtgr	5460
gcakggctct mttcctaktc tcagrcatc ccwgtggcam cctgtgcata gtctgaagac	5520
aatggaccca agtttaggtgt gacacgttca cataactgtg cagcacatgc cgggagtgtat	5580
cagtcagaca ttttaatttgc aaccacgtat ctctggtagt ctacaaagtt cctcaggat	5640
ttcattttgc aggcatgtct ctgagcttct atacctgctc aaggtcagtg tcatcttgc	5700
gttttagctca tccaaaggttgc ttaccctgggttcaatgaac ctaacctcat tctttgtatc	5760
ttcagtggttgc aattgtttta gctgatccat cttaacgcga ggagggatcc ttggctgagg	5820

95

attgtatttc agaaccacca actgctttg acaaattgtta acccgctagg ctcctttggt	5880
tagagaagcc acagtccccc agcctccaaat tggtgtcagt acttaggaag accacagcta	5940
gatggacaaa cagcattggg agaccttagc cctgctcetc tcgattccat cctgttagaga	6000
acaggaggta ggagccgctg gcaggagaca gcatgtcacc caggactctg ccgggtcaga	6060
atatgaacaa cgcctatgttc ttgcagaaaaa cgcttagcct gagtttcata ggaggtaatc	6120
accagacaac tgcagaatgt agaacactga gcaggacaac tgacctgtct cttcacata	6180
gtccatataca ccacaaatca cacaacaaaa aggagaagag atattttggg ttcaaaaaaaaa	6240
gtaaaaagat aatatacgat catttcttta gttattttga accccaaata tttccatc	6300
tttttgttgt tgtcattgtat ggtggtgaca tggacttgtt tatagaggac aggtcagctg	6360
tctggctcag tgcatacatat tctgaagtttgc tctgaaaatg tcttcatgtat taaattcagc	6420
ctaaacgttt tgccgggaac actgcagaga caatgctgtg agtttccaac cttagccat	6480
ctgcgggcag agaaggtcta gtttgcattc cagcattatc atgatatcag gactggttac	6540
ttgggttaagg aggggtctag gagatctgtc ctttttagag acaccttact tataatgaag	6600
tatattggag ggtggttttc aaaagttagaa atgtcctgtat ttccgatgtat catcctgtaa	6660
acatatttatac atttattaaat catccctgcc tgcgtctatt attatattca tatctctacg	6720
ctggaaactt tctgcctcaa tggactgtt gcctttgttt ttgcttagtt gtgttgtga	6780
aaaaaaaaac attctctgcc tgagttttaa ttttgtcca aagtttattt aatctataca	6840
attaaaaagct tttgcctcta gatcgccggc ggccgc	6876

<210> 73
<211> 3060
<212> DNA
<213> Homo sapien

<400> 73	
gcgtcgctga ggcccccattt gcttccccc gcccggcttgc ggcggggccca ctgtcgcccc	60
cgcgtcgccgg gcggcgccgg gtcgtcgctg gggccatggc tccggccggc cgcttcgtcc	120
tggagcttcc cgactgcacc ctggctcaat tggccctagg cggccacgccc cccggcgacg	180
cagacggcccc cggccccccgc ctggcgccggc tgctggggcc cccggagcgcc agctactcg	240
tgtgcgtgcc cgtgaccccg gacgcccggc gggggggcccg ggtccggggcg ggcggcgatgc	300
accagcgccct gctgcaccatg ctgcggccggc gcccatttcca gcggtgccag ctgtcgccgg	360
tgctctgcta ctggccgggc ggccaggcccg gcccggccaca gcaaggcttc ctgtcgcccg	420
accccgcttggaa tgaccctgac acccgccaaatg cgctgcgtcgat gctgctgggc ggcggccagg	480
aggcaccacg cccgcacttg ggcgagttcg aggccgaccc ggcggccag ctgtggcagc	540

96

gcctctggga	ggtgcaagac	ggcaggcgcc	tgcaggtggg	ctgcgcacag	gtcggtcccc	600
tcccgagcc	cccgctgcac	ccgggtggtgc	cagacttgcc	cagttccgtg	gtcttcccgg	660
acgggaagc	cgcggggcc	gtttggagg	agtgtacctc	ctttattcct	gaagcccggg	720
cagtgttga	cctggtcgac	cagtgccaa	aacagatcca	gaaaggaaag	ttccaggttg	780
ttgccatcga	aggactggat	gccacgggta	aaaccacggt	gaccagtc	gtggcagatt	840
cacttaaggc	tgtcctctta	aagtccaccac	cctcttgc	tggccagtgg	aggaagatct	900
ttgatgatga	accaactatc	attagaagag	ctttttactc	tttgggcaat	tatattgtgg	960
cctccgaaat	agctaaagaa	tctgccaaat	ctcctgtgat	tgttagacagg	tactggcaca	1020
gcacggccac	ctatgccata	gccactgagg	tgagtggggg	tctccagcac	ctggcccccag	1080
cccatcaccc	tgtgtaccag	tggccagagg	acctgctcaa	acctgacctt	atcctgctgc	1140
tcactgtgag	tcctgaggag	aggttgcaga	ggctgcaggg	ccggggcatg	gagaagacca	1200
gggaagaagc	agaacttgag	gccaacagt	tgttctgtca	aaaggtagaa	atgtcctacc	1260
agcggatgga	gaatcctggc	tgccatgtgg	ttgatgccag	cccctccaga	gaaaagggtcc	1320
tgcagacggt	attaagccta	atccagaata	gttttagtga	accgttagtt	ctctggccag	1380
gtgccacgtc	taactagatt	agatgttgtt	tgaaacatct	acatccacca	tttggtatgc	1440
agtgttccca	aatttctgtt	ctacaagcat	gttgtgtggc	agaaaaactgg	agaccaggca	1500
tcttaatttt	acttcagcca	tcgtaccctc	ttctgactga	tggaccgc	atcacaaagg	1560
tccctctcat	catgttccag	tgagaggcca	gcgattgctt	tcttcctggc	atagtaaaca	1620
ttttcttgg	acatatgttt	cacttaatca	ctaccaaata	tctggaagac	ctgtcttact	1680
cagacagcac	caggtgtaca	gaagcagcag	acaagatctt	ccagatcagc	agggagaccc	1740
cggagccct	gcttctcc	cactggcatg	ctgatgagat	cgtgacatgc	ccacattggc	1800
ttcttccaca	tctgggtgca	ctcgcatga	tgggctcgct	gcatctccct	cagtccaaa	1860
ttcttagagcc	aagtgttcct	gcagaggctg	tctatgtgtc	ctggctgccc	aaggacactc	1920
ctgcagagcc	atttttgggt	aaggaacact	tacaaagaag	gcattgtatct	tgtgtctgag	1980
gctcagagcc	cttttgatag	gcttctgagt	catatataaa	gacattcaag	ccaagatgct	2040
ccaactgcaa	atataccaac	cttctctgaa	ttatattttg	cttatttata	tttctttct	2100
tttttctaa	agtatggctc	tgaatagaat	gcacatttc	cattgaactg	gatgcatttc	2160
atttagccaa	tccagtaatt	tatttatatt	aatctataca	taatatgttt	cctcagcata	2220
ggagctatga	ttcattaatt	aaaagtggag	tcaaaacgct	aaatgcaatg	tttggtgtgt	2280
atttcatta	cacaaactta	atttgtcttg	ttaaataagt	acagtggatc	ttggagtggg	2340
atttcttgg	aaattatctt	gcacttgaat	gtctcatgat	tacatatgaa	atcgctttga	2400

catatcttta gacagaaaaa agtagctgag tgaggggaa attatagagc tgtgtactt	2460
tagggagtag gttgaaccag gtgattacct aaaattcctt ccagttcaaa ggcagataaa	2520
tctgtaaatt attttacccat atctaccatt tcttaagaag acattactcc aaaataatta	2580
aatttaaggc tttatcaggt ctgcatatag aatcttaaat tctaataaag tttcatgtta	2640
atgtcatagg atttttaaaa gagctatagg taatttctgt ataatatgtg tatattaaaa	2700
tgtaattgat ttcatgtaa agtatttaa agctgataaa tagcattagg gttcttgca	2760
atgtggtatac tagctgtatt attggttta ttactttaa acatttgaa aagcttatac	2820
tggcagccctaa gaaaaacaaa caattaatgt atctttatgt ccctggcaca tgaataaact	2880
ttgcgtggt ttactaatct atgctgtcat cctgggtaca tattgatttgc tctaaaaagt	2940
gcttctcag attcccctt taatattgtg atgtaaagga gggaaatttt ggtaaaggaa	3000
gttgaaggt gtgagctggc aggctaagtg gaatttgc ttcaagtgtct ttcaagagaaa	3060

<210> 74
<211> 3885
<212> DNA
<213> Homo sapien

<400> 74	
gaaaagtact ggaagtaaag tctgaccaa agcaaatgaa cagcttaacc ggagatcaca	60
aaggctacaa caattaacag aggtttcaag aaggtcgtta cgcaatcgaa aaattcaggg	120
tcaagttcaa gcagttaaac agagtttgcc accaactaaa aaagagcagt gtagcgtac	180
tcagagtaaa tctaataaaaa caagtcaaaa acatgtgaag agaaaagtac tggaaagtaaa	240
gtctgactct aaagaagatg aaaatctgt aattaatgaa gtaataaatt ctcccaaagg	300
gaaaaaaaaacgc aaggtagaac atcagacagc ttgtgcttgt agttctcaat gcatgcaagg	360
atctgaaaag tgcctcaga agactactag aagagacgaa acgaaacctg tgcctgtac	420
ttctgaggtg aaaagatcaa aaatggctac ttcaagtggc ccgaaaaaga atgagatgaa	480
gaagtcgggt catacacaag tgaataactaa cacaacactc ccaaaaagtc cacagccatc	540
agtgcctgaa caaagtgata atgagctgga gcaagcagga aagagcaaac gaggtgtat	600
tctccagctc tgtgaagaaa ttgctggtga aattgagtca gataatgtag aggtaaaaaa	660
ggaatcttca caaatggaaa gtgtaaagga agaaaagccc acagaaataa aatttggaaaga	720
gaccagtgtt gaaagacaaa tacttcatca gaaggaaaca aatcaggatg tgcaatgtaa	780
tcgttttttc ccaagtagaa aaacaaagcc tgtgaaatgt atactaaatg gaataaacag	840
ctcagccaaag aagaactcca actggactaa aatcaaactc tcaaaaatttactctgtgca	900
gcacaataag ttggactctc aagttcccc taaatttaggc ttattacgaa ccagtttttc	960

accaccagct ttagaaatgc atcatccagt gactcaaagt acattttag ggacaaagct	1020
acatgataga aatataactt gccagcagga aaaaatgaaa gaaattaatt ctgaagaagt	1080
aaaaattaat gatattacag tagaaattaa taaaaccaca gaaagggttc ctgaaaattg	1140
tcatttggcc aatgagataa aaccttctga cccaccattg gataatcaga tgaaacattc	1200
tttgattca gcatcaaata agaatttcag ccaatgttg gaatccaagc tagaaaacag	1260
tccagtgaa aatgttactg ctgcttcgac tctgctcagt caagaaaaaa ttgatacagg	1320
agagaataaa tttccaggtt cagctccccca acagcatagt attctcagta accagacatc	1380
taaaagcagt gataacaggg agacaccacg aaatcattct ttgcctaagt gtaattccca	1440
tttggagata acaattccaa aggacttgaa actaaaagaa gcagagaaaa ctgatgaaaa	1500
acagttgatt atagatgcag gacaaaaaaag atttggagca gtttcttgcata atgtttgtgg	1560
aatgctgtat acagcttcaa atccagaaga tgaaacacag catctgctt tccacaacca	1620
gtttataagt gctgttaat atgtggttct gctcattaat caccacgagt gtggatctga	1680
agaagagttt attacctctc ttttttgag tatgttaac ttcagataca cacaacgtag	1740
cttctccttc cctattagat tcttagaagg gctggaagaa agaaagaatt ctggctgaat	1800
accctgatgg caggataata atggttcttc ctgaagaccc aaagtatgcc ctgaaaaagg	1860
ttgacgagat tagagagatg gttgacaatg atttaggtt tcaacagggct ccactaatgt	1920
gctattccag aactaaaaca cttctttca tttccatga caaaaaagta gttggctgcc	1980
taattgcgga acatatccaa tggggctaca gagttataga agagaaactt ccagttatca	2040
ggtcagaaga agaaaaagtc agatttgaaa ggcaaaaagc ctgggtctgc tcaacattac	2100
cagagcctgc aatctgcggg atcagtcgaa tatgggtatt cagcatgatg cgtcggaaga	2160
aaattgcttc tcgcatttgcattt gaatgcctaa ggagtaactt tatatatggc tcataattga	2220
gcaaaagaaga aattgcttcc tcagatccca ctcctgatgg aaagctgtt gcaacacagt	2280
actgtggcac tggcaattt ctggtatata attttatcaa tggacagaat agcacgtaaa	2340
acaaattctt gcctacacca cttagaagaca tctattgaag agaatggatt ggttgcgtac	2400
ttaaccagg aactagggcc atttttatttta caatgaactc aggactggca acaaccatat	2460
ggttgcgttcca ttttcataaaa attggaaaca atgcagtaat agcttattgt tttgtttttt	2520
aaagaagata ttttatttac ttttacagaa atttatgatt gatgtatttt atctatagtt	2580
atttagacat gtttacatgc agcagataat ttttcataatg ggactgaaaa ctaatgcaag	2640
gactatggtc tcagtgataa gtatatttg aagttcttaa tatggaaata taccagtgt	2700
gcttggtact gtatttttt atattgatct gctgataccca gtgataggct taaagattgt	2760

99

attttcacag	agtggaaacc	aatttttta	gttattgttc	aaggagggtg	caatattaag	2820
tgtttggaa	tttgaagcta	attttaaaa	ggcctgaact	atactttgaa	gaaaccctta	2880
tagaaaagga	aagctccagc	taaataggaa	gaattagaat	attgagctt	ttttcctga	2940
tttttctctt	tcctatcttt	gatggaagga	ggaagtagaa	agtggtaaag	aattgaggct	3000
ttccttcctt	gagagctgta	aatgacaagc	attaggaaag	gtaccctcct	agattcatta	3060
ttcttcatt	ctggttcac	ttttaaaaata	aatggcaact	tggcacacct	aggctgttaa	3120
caaactctcaa	agaggttat	aaaaacgtat	agaatacttg	gaagcaaagt	atggatgact	3180
cggtatctgc	tttggatttc	ctcagaaata	ctgcactgag	tatatgccct	cattactgga	3240
cttcattttg	atacttgtct	atccttcata	gtgccctcta	cttttaaagg	gtttatatgt	3300
tgaaaaactg	ctgtggcctt	ttatgacctg	tatataatgt	agaataaaaa	taataaaaata	3360
cttgatagct	ttttctaagt	gaccaatgta	ctaaactgaga	ataatggtgt	gttgcattt	3420
gtgcttttc	agggtgtttt	tttggttga	tatctgaaa	tatgattaaa	acattggctt	3480
cctaaaggca	gtttccacca	gttgccaaa	ggatcattgt	gtcagcagca	aatcagctga	3540
actttatttc	caaaggcaaa	atccttctg	attattttag	taacatagta	cttttatgat	3600
gttgcaaata	aatgaagggc	ccacagccca	agaatgaatt	accactgtgg	ttcaacttag	3660
gttattttg	tgagctgaaa	tgatcatatc	tcagttgaaa	actggctaaa	attttagggcc	3720
ttaaattaac	aggtatacat	tttatttccc	tataaatttt	tgctttaca	atttcttaggc	3780
cactgcacct	ggccctagct	tttgatactg	tcatttcct	ttgggcttga	gactgttcta	3840
gtcaatcctg	gtctcattgt	ttgcctgaca	ggtaccatga	tttaa		3885

<210> 75
<211> 2271
<212> DNA
<213> Homo sapien

<400> 75						
aggatgatag	atatataggc	gaatggkctc	tagatcatgc	tcgagcggcg	cagtgtatg	60
gatgcgttgt	cgcggccgta	cagcgtggag	tggatggct	ctcttccctc	agccacgccc	120
cttgcgtgagga	cagaggtggg	ggagtggaa	gtggaaagtc	accagagaac	aggagaggga	180
tttgaggggcg	cgacccccagc	gctctccacg	gaccagccag	agggactgga	gccaggtgtg	240
catgggttca	aggccctggc	cctgcccagc	ctctgtcttg	ggagctcagc	cccagggttc	300
ggtcgtcagc	agtttcccaa	gaacaagatg	tgatggcatc	tgctgctgaa	accctgatga	360
ggaccaggcc	ccctgcaccc	ctgtcagcct	gaggaattaa	agctttggtg	ctgggaagag	420
cattattcct	ctgaggagcc	gctgtgcctc	tttctgaagt	gagggccgtg	ccccgggtcc	480

100	
catttctcct ttcacttgag tcgggaagca cagcaacttt aaggctcgcg cccagcaaca	540
tggctccccct cgcatctgca tctccctcct gctctggtgt tgccgctgca ccctgtcctc	600
ggaggacagc agaggtttgg acggagactc agggagggag ggaaggaggc aaggacgcct	660
gtggaaacat ctttcaggca gctcttagggt ctgggggcca ggatgcctgg gtctcccaag	720
gcctgtctgc tgtctctgcc accctcagcg gctgccagaa gcagcgtgtg ggggaggcat	780
gtgctgcagc acacctgcgg ccgagaccag cactcagagg tggctcccc tgacaggaac	840
cgtgttagggt gcagaaggct gagacctgtg gacactgcgt gtttatggc agcttgcttg	900
ctggggctca tggccacagt ggagagggc cgtgggtcag ggcagccccc tgtgcagtcc	960
agtgcgggc aggagtcttgc caggggctca tgaccacagt ggagagggc tgtgggtcag	1020
gggcagcctg gcgtgcagtc cagtgcggg caggagtctc acaggggctc gtggccacag	1080
tggagagggc ctgtgggtca gggggcagcc cggcatgcag tccagtgcgg ggcaggagtc	1140
tcgcagaatg cagcctgacg cctccacgtg gctccccgg cccctacagg ctccctcagc	1200
tgcagagctg ggtcccatcc gacgctgtcg ctgggcagcg agaggcagag gcaggttccc	1260
cgagggaaagc atggggccct tctccggcc acggttgcac cagcaggagt tcattttgc	1320
agccccagag ccagggtgat gtgggcacag gtgtcaagtc agggtggtcg ttagccttgc	1380
gcccccagga gagatatggc ctgaagcctg ctgcacgtgc gtgccacacg cgtgtggggc	1440
cacctctgca catcctgagg tgaccctttt ggggggtcg ttaggtcag tgcacgtgtg	1500
ccggcagggc tggtcagggt tcattgcctg cccaggagcc tgagcctgag gcagggaggt	1560
gctggtgacc gttcccaa ggtggctcac ccacagcacc gggaatggac caggtcgcc	1620
ctgcccctca gtaagcctgg ggactggcag accgtcttt ttctggggac acgtatccag	1680
ccacacatgg gctgacccccc tcccagtctc tgaccccgac acagtttgcat cccttctcag	1740
gccaatcctg aggctcaggg ctggcacact gtctctatcc caaggcaagc acaggtgggc	1800
acactgcctt tgcctttgtt ccactgtggg actggcctg tctgtctcca gcgcggcagca	1860
tggcctccac acacctctgc ctccaggcgt ggctgggcct gcccctcagag tccctgccac	1920
gccagccgtt ggctgcaggc atatcacaga tagggatgc tgcccaggc tccgagtaga	1980
ccaaaagatt cctgcccaca gcccaggaag agcaggcagg caacggcgat tccccggaa	2040
gggaaggggcc ccggagtgaaa gtgtcagaa ccctgggcca ctgtgctgtt aaccaccacc	2100
tcccggaat ggctggcctc agcgaggccc cagggcctcc cccgcggcctc gcagtgtgca	2160
tgtccctggc cctctccat caccaggctg tggtggtgt gtggggaggc tgtggtacac	2220
aacgcaggtaa aaataatatg agaacatgca cccagcacca ggggactcag a	2271

101

<210> 76
 <211> 2186
 <212> DNA
 <213> Homo sapien

<400> 76
 aggatgatag atatataggc gaatggkctc tagatcatgc tcgagcggcg cagtgtatg 60
 gatgcgttgt cgccggccgcg gcccgcggg caggtcgca gggcgccacc ccagcgctct 120
 ccacggacca gccagaggga ctggagccag gtgtgcattt gttcaaggcc ctggccctgc 180
 ccagcctctg tcttgggagc tcagccccag gtttcggtgc tcagcagttt cccaagaaca 240
 agatgtatg gcatctgctg ctgaaaccct gatgaggacc aggcccccgt caccgcgtgc 300
 agcctgagga attaaagctt tggtgctggg aagagcatta ttccctctgag gagccgctgt 360
 gtttccttctt gaagtgaggg ccgtgcggccg ggtcccatcc ttcccttcac ttgagtcggg 420
 aagcacagca actttaaggc tcgcgcggccag caacatggct cccctcgcat ctgcatttcc 480
 ctccctgtctt ggtgttgcgg ctgcaccctg tcctcggagg acagcagagg tttggacgga 540
 gactcaggga gggagggaag gaggcaagga cgcctgtgg aacatcttcc aggccagctct 600
 agggtctggg ggccaggatg cctgggtctc ccaaggcctg tctgctgtct ctgccaccct 660
 cagcggctgc cagaaggcagc gtgtggggga ggcattgtgt gcagcacacc tgccggccgag 720
 accagcactc agaggtcgcc tcccctgaca ggaaccgtgt agggtgcaga aggctgagac 780
 ctgtggacac tgcgtgtttt atggcagctt gcttgcattt gctcatggcc acagtggaga 840
 gggggccgtgg gtcagggcag cccgggtgtc agtccagtgc cgggcaggag tcttgcagg 900
 gctcatgacc acagtggaga ggggctgtgg gtcagggca gcctggcgtg cagtccagtgc 960
 cccggcagga gtctcacagg ggctcggtgc cacagtggag aggggctgtg gtcagggggg 1020
 cagcccccga tgcagtccag tgccggcag gagtctcgca gaatgcagcc tgacgcctcc 1080
 acgtggctcc cccggccccc acaggctccc tcagctgcag agctgggtcc catccgacgc 1140
 tgtcgctggg cagcgagagg cagaggcagg ttcccccagg gaagcatggg ccccttctcc 1200
 cggccacggc tgccccagca ggagttcatc tttgcagccc cagagccagg gtgtatgtgg 1260
 cacaggtgtc aagtcaagggt ggtcggttagc cttgcgcggc caggagagat atggcctgaa 1320
 gcctgctgca cgtgcgtgcc acacgcgtgt gggccaccc ctgcacatcc tgaggtgacc 1380
 cttttgggggg ggtcgatg gtcagtgcac gtgtgcggc agggctggc agggttcatc 1440
 gcctgcccag gagcctgagc ctgaggcagg gaggtgtgg tgaccgttcc cccaagggtgg 1500
 ctcacccaca gcacccggaa tggaccaggt cgtccctgca cctcagtaag cctggggact 1560
 ggcagaccgt ctctttctg gggacacgta tccagccaca catgggctga cccctccca 1620
 gtctctgcac ccgacacagt ttgatccctt ctcaggccaa tcctgaggct caggcgtggc 1680

102

acactgtctc tatcccaagg caagcacagg tggcacact gcccgttgc ttggtccact	1740
gtgggactgg tcctgtctgt ctccagcgcc cagcatggcc tccacacacc tctgcctcca	1800
gggctggctg ggccctgccct cagagtccct gccacgccag ccgttggctg cagggatatac	1860
acagataggg gatgctgccc agggctccga gtagacccaa agattcctgc ccacagccca	1920
ggaagagcag gcaggcaacg gcgattcccc gggaaaggaa gggcccccggaa gtggggtgct	1980
cagaaccctg ggccactgtg ctgttaacca ccacctcccg gcaatggctg gcctcagcga	2040
ggccccaggc cctcccccgc gcctcgcagt gtgcattgtcc ctggccctct cccatcacca	2100
ggctgtggtg ggtgtgtggg gaggctgtgg tacacaacgc aggtaaaata atatgagaac	2160
atgcacccag caccagggga ctcaga	2186

<210> 77
<211> 1258
<212> DNA
<213> Homo sapien

<400> 77	
tgatggatcg gccgccccggg caggtcaaag cggcaacaag tgatctggaa cactatgaca	60
agactcgtca tgaagaattt aaaaaatatg aaatgtgaa ggaacatgaa aggagagaat	120
atttaaaaac attgaatgaa gaaaagagaa aagaagaaga gtctaaattt gaagaaatga	180
aaaaaaagca tgaaaatcac cctaaagtta atcacccagg aagcaaagat caactaaaag	240
aggatggaa agagactgat ggattggatc ctaatgactt tgaccccaag acattttca	300
aattacatga tgtcaatagt gatggattcc tggatgaaca agaatttagaa gccctattta	360
ctaaagagtt ggagaaagta tatgacccta aaaatgaaga ggtatgatatg gtagaaatgg	420
aagaagaaag gcttagaatg agggAACATG taatgaatga ggTTgatact aacaaagaca	480
gattggtgac tctggaggag tttttgaaag ccacagaaaa aaaagaattc ttggagccag	540
atagctggaa gacatttagat cagcaacagt tcttcacaga ggaagaacta aaagaatatg	600
aaaatattat tgctttacaa gaaaatgaac ttaagaagaa ggcagatgag ctccagaaac	660
aaaaagaaga gctacaacgt cagcatgatc aactggaggc tcagaagctg gaatatcatc	720
aggcatacata gcagatggaa caaaaaaaaaat tacaacaagg aattcctcca tcagggccag	780
ctggagaatt gaagtttgag ccacacattt aaagtctgaa gtccaccaga acttggaaaga	840
aagctgttaa ctcaacatct atttcatttttttttcc ttccttttc tctgctcaat	900
aaatattttta aaagcatatt tggaaataaag ggagatactt tttaaatgaa aacacttttt	960
ttgggacaca gatattaaag gattgaagtt tatcagaacc aggaagaaaa caaactcact	1020
gtctgtctc tgctctcaca ttcacacggc tcttttattt atttttttgt tctcccttaa	1080

tgatttaatt aagtggcttt atgccataat ttagtgaaac tattaggaac tatttaagtg	1140
agaaaaactct gcctcttgc tttaaattag attgctctca cttactcgta aacataggta	1200
ttcttttatg ggtgcttatac attccttctt tcaataaatg tctgtttgat attaaca	1258

<210> 78
<211> 1597
<212> DNA
<213> Homo sapien

<400> 78	
gaagagggtg ataaaaggaaa ggagaaggcc attcttactg acctgatagt ggaagaaaaa	60
tgagggtggag gaccatcctg ctacagtatt gcttcttctt gattacatgt ttacttactg	120
ctcttgaagc tgcgcattt gacatagaca agacaaaagt acaaaaatatt caccctgtgg	180
aaagtgcgaa gatagaacca ccagatactg gactttatta tgcataatat ctcaagcaag	240
tgattgtgt gctggaaaca gataaacact tcagagaaaa gctccagaaaa gcagacatag	300
aggaaaataaa gagtgggagg ctaagcaaag aactggattt agtaagtccat catgtgagga	360
caaaaacttga tgaactgaaa aggcaagaag taggaaggaa aagaatgtt attaaagcta	420
atggatttc cttcaagat ataggcatgg accaccaagc tcttctaaaa caatttgatc	480
acctaaacca cctgaatcct gacaaggttt aatccacaga ttttagatatg ctaatcaaag	540
cgccaacaag tgcgcattt gactatgaca agactcgta tgaagaattt aaaaaatatg	600
aaatgtgaa ggaacatgaa aggagagaat attaaaaac attgaatgaa gaaaagagaa	660
aagaagaaga gtctaaattt gaagaaatga agaaaaagca tgaaaatcac cctaaagttt	720
atcacccagg aagcaaagat caactaaaag aggtatgggaa agagactgtat ggattggatc	780
ctaattgactt tgaccccaag acattttca aattacatga tgtcaatagt gatggattcc	840
tggatgaaca agaatttagaa gccctatttta ctaaagagtt ggagaaagta tatgacccta	900
aaaatgaaga ggatgatatg gtagaaatgg aagaagaaaag gcttagaatg agggAACATG	960
taatgtgaa ggttataact aacaaagaca gattgggtac tctggaggag tttttgaaag	1020
ccacagaaaaaaa aaaagaatttccat tggagccag atagctgggaa ggtcatacag cagatggAAC	1080
aaaaaaaaattt acaacaagga attcctccat cagggccagc tggagaattt aagtttgc	1140
cacacatttta aagtctgaaatccat tccaccagaa ctttggaaagaa agctgttaac tcaacatcta	1200
tttcatcttt ttagctccct tccttttct ctgctcaata aatattttaa aagcatattt	1260
gaaataaagg gagatactttt ttaaatgaaa acactttttt tgggacacag atattaaagg	1320
attgaagttt atcagaacca ggaagaaaac aaactcactg tctgctctt gcttcacat	1380
tcacacggct cttttattttt tttttttgtt ctccttaat gatttaatttta agtggcttta	1440

104

tgccataatt tagtgaaact attaggaact atttaagtga gaaaactctg cctcttgctt	1500
ttaaattaga ttgctctcac ttactcgtaa acataggtat tctttatgg gtgcttatca	1560
ttccttcctt caataaatgt ctgtttgata ttaacaa	1597

<210> 79
<211> 1959
<212> DNA
<213> Homo sapien

<400> 79	
ggggcagagc ggagcggtgg gccggggct ggaggacagg tttgtgcgct ggacgcaagc	60
accaggcgca gcctcgctcg ccgacaccccg gccagaacgt gttacgagtc agtttttagt	120
aaaaaaacat tgagcttagga gccaagaccc atctcttcac tattttggta ttgtgcaagt	180
catcttacct ctctggatct cagttgtctc atctgtaaaa aggagataaa aattatttac	240
ctgcctgaac atgaggtgga ggaccatcct gctacagtat tgctttctct tgattacatg	300
tttacttaact gctcttgaag ctgtgcctat tgacatagac aagacaaaaag tacaaaatat	360
tcaccctgtg gaaaagtgcga agatagaacc accagatact ggactttatt atgatgaata	420
tctcaagcaa gtgattgatg tgctggaaac agataaacac ttcagagaaaa agctccagaa	480
agcagacata gaggaaataa agagtggag gctaagcaaa gaactggatt tagtaagtca	540
ccatgtgagg acaaaaacttg atgaactgaa aaggcaagaa gtaggaaggt taagaatgtt	600
aattaaagct aagttggatt cccttcaaga tataggcatg gaccaccaag ctcttctaaa	660
acaatttgat cacctaaacc acctgaatcc tgacaagttt gaatccacag atttagat	720
gctaatcaaa gcggcaacaa gtgatctgga acactatgac aagactcgtc atgaagaatt	780
taaaaaatata gaaatgatga aggaacatga aaggagagaa tattttaaaaa cattgaatga	840
agaaaaagaga aaagaagaag agtctaaatt tgaagaaatg aagaaaaagc atgaaaatca	900
cccttaaagtt aatcacccag gaagcaaaga tcaactaaaa gaggtatggg aagagactga	960
tggattggat cctaatgact ttgacccaa gacatttttc aaattacatg atgtcaatag	1020
tgtatggattc ctggatgaac aagaattaga agccctatTTT actaaagagt tggagaaagt	1080
atatgaccct aaaaatgaag aggatgatATG ggtagaaatg gaagaagaaaa ggcttagaat	1140
gagggAACAT gtaatgaatg aggttgatac taacaaAGAC agattggtga ctctggagGA	1200
gtttttgaaa gccacagaaa aaaaagaatt ctggagcca gatagctggg agacattAGA	1260
tcaGCAACAG ttcttcacAG aggaAGAACT aaaAGAATAT gaaaATATTa ttgctttaca	1320
agaaaaatgaa cttaAGAAGA aggCAGATGA GCTTCAGAAA caaaaAGAAG agctacaACG	1380
tcaGCAATGAT caACTGGAGG CTCAGAAGCT GGAATATCAT CAGGTCAATAC AGCAGATGGA	1440

105

acaaaaaaaaaa ttacaacaag gaattcctcc atcagggcc a gctggagaat tgaagtttga	1500
gccacacatt taaagtctga agtccaccag aacttggaaag aaagctgtta actcaacatc	1560
tatttcatct ttttagctcc ctccctttt ctctgctcaa taaatatttt aaaagcatat	1620
ttgaaataaa gggagatact tttaaatga aaacactttt tttgggacac agatattaaa	1680
ggattgaagt ttatcagaac caggaagaaa acaaactcac tgtctgctct ctgctctcac	1740
attcacacgg ctctttatt tatttttg ttctccttta atgatttaat taagtggcct	1800
tatgccataa ttttagtgaaa ctattaggaa ctatthaagt gagaaaactc tgccttgc	1860
ttttaaatta gattgctccc acttactcgt aaacataggt attctttat gggtgcttat	1920
cattccttct ttcaataaaat gtctgttga tattaacaa	1959

<210> 80
<211> 1625
<212> DNA
<213> Homo sapien

<400> 80	
aaaaagcaaa gagtaccaga ctcacaagta tggttatgag agctacatga tatagtatat	60
agcaaaggaa tttatttagtt taaaagtact atggaaatgt taattttggaa aatgtgaggt	120
aatatttata aggcaacttag aacaatgcta gccacatagt gtttggtaaa tagattaaaa	180
cagtcctagt aatatcgta tcttagaata cacagttcat gttattgcac caaagctact	240
tctgaaatga ctaaagatag ccacttggtt caatatacct gagaaaatag agtgaagtt	300
ttataaaaaaaaa tgtagtctg taatgcaaac ttcaagtact tgggaaatcc cttccccac	360
aaacagttta gtagtgaagt tgcactctat ggacaaaatt acctactatc acaaataaaa	420
aaagtgtata ttcaagcgctc tgagggccag aaatactcgg agatcaatta aactagatgg	480
aaaaggagaa cccaaagggg cgaagagagc gaagccagtg aagtacactg cagcaaagct	540
gcatgagaaa ggtgtcctgc tagatataga tgatctcaa acaaaccagt ttaagaatgt	600
tacatttgat atcatagcta ctgaagatgt aggcatatc gatgttaagat caaaattcct	660
tggtgtttagt atggaaaagg tgcaactcaa tattcaggat ttacttcaga tgcaatata	720
aggatgtact gtaatgaaaaa tgtttgataa ggttaaagtg aatgtaaacc ttctcatata	780
cctgctgaac aagaagttct atggaaagtg aagtgcctac agaaatttct tggattctgt	840
atcatctgga ttaggaaatg aatttggtaa atattttgc ttttaaacat gattgaaatc	900
actgcttata aatgtgtgat ttttttaaa cgacaaaaac tggctgaaag aatgtaccca	960
ggtgccttt tgctaatttg atactataat agaatgagac ataaaaatgaa ttaatggaaa	1020
catatccaca ctgtactgtg atataggtac tctgattaa aactttggac atcctgtgat	1080

106

ctgttttaaa gttgggggggt gggaaattta gctgactagg gacaaacatg taaaccttatt	1140
ttcctatgaa aaaaatttta aatgtcccac ttgaataacg taattcttca tagttttttt	1200
aatctatgga taaatggaaa cctaattatt tgtaatgaat tattnagaca gttctaagcc	1260
ctgtcttctg ggagtttatca attttaaaga gaactttgt gcaattcaaa tgaagttttt	1320
ataagtaatt gaaaatgaca acacaataac actttctgta taaaagtata tattttatgt	1380
gatttattcc tactaatga aagtgcacta ctgcctcatg taaagactct tgcacgcaga	1440
gcctttaagt gactaaggaa caacatagat agtgagcata gtccccacct ccaccctca	1500
caatttatcc gaataacttca attgtgcctc tcaattttt gtaatgctaa aaaatcagta	1560
tctagatggt tttaaatgt attctctgga aattgtttta tgtaaaataa atgttactta	1620
attcc	1625

<210> 81
<211> 772
<212> DNA
<213> *Homo sapien*

<400> 81
gcaaaggcagc gcggaaagcag gggggggcga cgagcgagaa attaacacgt atgggcgatg 60
ggccctaattt caatgcgagc ggccgcagtgt gatggatgtc cgccggcgagg tacttctgag 120
ctgccttaat gcaaggcat ttatatttgt taagaggaaa taatcaagat cactcatatc 180
ccaactgaat ctgaggtttt ataaatccct caaacgattt ctgagagcct gattgtggaa 240
agaagtgaga tgcacccattt tttcaagaag tcctggaaag cgctctccctt gcacgtccat 300
ttccaggagg agaagcaagc agatgagagg tttccattt tgtcatccaa ggttagctgtg 360
cacttgcctt gttgctgaag ttccaataat gtgaaaacca aagtagaggt tttttcttc 420
ttctttttgtt tttctattaa ttccacttat accaaagtgt ttgaaagttt gaaatgtgtt 480
gtttctgagt tatataaggc tacttcatga caagactgct ttgtaatatt tcactttgtt 540
ttactacaaa ttccatcac tttgttttac tataaattca gattatccaa atatttccct 600
aataactatgtt gggaaatgctg attttctttt gttacgttgtt gggaaacattt tgcatgttt 660
acatagttctt catgaaacat gggaaatttttt gaaagtgtata tatgatacac attttttggtt 720
tatgttattttt aattttgtgtt aataaaggcag taacatttttgcatttttta ag 772

<210> 82
<211> 3198
<212> DNA
<213> *Homo sapien*

<400> 82

107

ggcactggcc ttccatggca cagcacccct gccaactgg cgctggctgg tctacgacaa	60
gctcagcccc atccccaaaca acaacggctt catcaaccag gacttcgtgg tgtggatgcg	120
catggcagcg ctgcccacgt tccgcaagct gtccgcaag ctgtacgggc acatccgcca	180
gggcaactac tcagctggc tgccgcggtg tgtctactgt gtcaacatca cctacaacta	240
cctggtaaga agcgcaattc cacactctac ataaccatgt tactcattgt tccagtcattc	300
gtcgcaggta caatcatagt actcctgctt tacctaaaa ggctcaagat tattatattc	360
cctccaaattc ctgatcctgg caagatttt aaagaaatgt ttggagacca gaatgatgat	420
actctgcact ggaagaagta cgacatctat gagaagcaaa ccaaggagga aaccgactct	480
gtatgtctga tagaaaacct gaagaaagcc tctcagtatggagataatt tattttacc	540
ttcactgtga ccttgagaag attcttccca ttctccattt gttatctggg aacttattaa	600
atggaaactg aaactactgc accattttaaa aacaggcagc tcataagagc cacaggttt	660
tatgttgagt cgccgcaccga aaaactaaaa ataatggcgttggagaa gagtgtggag	720
tcattctcat tgaattataa aagccagcag gcttcaaact aggggacaaa gcaaaaagtg	780
atgatagtgg tggagttaat cttatcaaga gtttgacaa ctccctgagg gatctatact	840
tgctttgtgt tctttgtgtc aacatgaaca aattttatgtt gtaggggaaac tcatttgggg	900
tgcaaatgct aatgtcaaacc ttgagtcaca aagaacatgt agaaaacaaa atggataaaaa	960
tctgatatgt attgtttggg atcctattga accatgtttg tggctattaa aactcttttta	1020
acagtctggg ctgggtccgg tggctcacgc ctgtatccc agcaatttgg gagtccgagg	1080
cgggcggatc actcgaggc aggagttcca gaccagcctg accaaaatgg tgaaacctcc	1140
tctctactaa aactacaaaa attaactggg tgggtggcg cgtgcctgtatcccagcta	1200
ctcgaaaagc tgaggcaggt gaattgttttgc aacctggag gtggaggttgc cagtgagcag	1260
agatcacacc actgcactct agcctgggtg acagagcaag actctgtcta aaaaacaaaa	1320
caaaaacaaaa caaaaacaaaa aaacctctta atattctggatcatcatc cttcgacag	1380
cattttcctc tgctttgaaa gccccagaaaa tcagtgtgg ccatgatgac aactacagaa	1440
aaaccagagg cagcttctt gccaagaccc ttcaaagcca ttttaggctg ttagggcag	1500
tggaggtaga atgactcctt gggatttgc gtttcaacca tgaagtctct aacaatgttat	1560
tttcttcacc tctgctactc aagtagcatt tactgtgtct ttggtttgc ttagggcccc	1620
gggtgtgaag cacagacccc ttccagggttacagtcta ttgagactc ctcagtttttctt	1680
gccacttttt ttttaatct ccaccagtca ttttcagac ctttaactc ctcaattcca	1740
acactgattt cccctttgc attctccctc cttcccttcc ttgtagcctt ttgactttca	1800
ttggaaatta ggtgttaaat ctgctcagga gacctggagg agcagaggat aatttagcatc	1860

tcaggtaag	tgtgagtaat	ctgagaaaca	atgactaatt	cttgcatttt	ttgttaacttc	1920
catgtgaggg	ttttcagcat	tgatatttgt	gcattttcta	aacagagatg	agggtggatc	1980
ttcacgtaga	acattggtat	tcgcttgaga	aaaaaaaagaat	agttgaacct	atttctcttt	2040
ctttacaaga	tgggtccagg	attcctcttt	tctctgccat	aatgattaa	ttaaatagct	2100
tttgtgtctt	acattggtag	ccagccagcc	aaggctctgt	ttatgctttt	ggggggcata	2160
tattgggttc	cattctcacc	tatccacaca	acatatccgt	atatatcccc	tctactctta	2220
cttcccccaa	atttaaagaa	gtatggaaa	tgagaggcat	ttccccccacc	ccatttctct	2280
cctcacacac	agactcatat	tactggtagg	aacttgagaa	ctttatttcc	aagttgttca	2340
aacatttacc	aatcatattta	atacaatgat	gctatttgca	attcctgctc	ctaggggagg	2400
ggagataaga	aaccctca	ctctacaggt	ttgggtacaa	gtggcaacct	gcttccatgg	2460
ccgtgttagaa	gcatggtgcc	ctggcttctc	tgaggaagct	ggggttcatg	acaatggcag	2520
atgtaaagtt	attcttgaag	tcagattgag	gctggagac	agccgtagta	gatgttctac	2580
tttgttctgc	tgttctctag	aaagaatatt	tggttttcct	gtataggaat	gagattaatt	2640
cctttccagg	tatTTATAA	ttctggaaag	caaaacccat	gcctccccct	agccattttt	2700
actgttatcc	tatTTAGATG	gccatgaaga	ggatgctgtg	aaattcccaa	caaacattga	2760
tgctgacagt	catgcagtct	gggagtgccc	aagtgtatctt	ttgttccat	cctcttcttt	2820
tagcagtaaa	atagctgagg	gaaaagggag	ggaaaaggaa	gttatggaa	tacctgtggt	2880
ggttgtgate	cctaggcttt	gggagctctt	ggaggtgtct	gtatcagtgg	attttccatc	2940
ccctgtggga	aatttagtagg	ctcatttact	gttttaggtc	tagcctatgt	ggatttttc	3000
ctaacataacc	taagcaaacc	cagtgtcagg	atggtaattc	ttattcttcc	gttcagttaa	3060
gttttccct	tcatctgggc	actgaaggga	tatgtgaaac	aatgttaaca	tttttggtag	3120
tcttcaacca	gggattgttt	ctgttaact	tcttataagga	aagcttgagt	aaaataaaata	3180
ttgtctttt	gtatgtca					3198

<210> 83
<211> 5193
<212> DNA
<213> Homo sapien

<400> 83						
cgggctgcag	gaattggcac	gaggagcgcg	acacatcctg	gagctggcgg	gcccgcagc	60
aaatgggacc	aaccagctcc	agccccactt	ctttcctcc	cgccagcggc	cccaggtggg	120
gaggtcacca	gcagtggggg	aagtccctggg	ggcaccacag	ctgctccttc	aggagccttgc	180
gatgctgctg	ctgctgtggc	tgccaagatt	aatgccatgc	tcatggcaaa	agggaaagctg	240

aaaccaactc agaatgcttc tgagaagctt caggctcctg gcaaaggcct aactagcaat	300
aaaagcaagg atgacctggt gtagctgaa gtagaaattta atgatgtgcc tctcacatgt	360
aggaacttgc tgactcgagg acagactcaa gacgagatca gccgacttag tggggctgca	420
gtatcaactc gagggagggtt catgacaact gaggaaaaag ccaaagtggg accaggggat	480
cgtccattat atcttcatgt tcagggccag acacggaaat tagtggacag agctgtaaac	540
cggatcaaag aaattatcac caatggagtg gttcaccagc cagcacccat cgctcagttg	600
tctccagctg ttagccagaa gcctcccttc cagtcaggga tgcattatgt tcaagataaa	660
ttatTTgtgg gtctagaaca tgctgtaccc acttttaatg tcaaggagaa ggtggaaggt	720
ccagggctgct cctatttgca gcacattcag attgaaacag gtgc当地 cttcctgcgg	780
ggcaaagggtt caggctgcat tgagccagca tctggccgag aagctttga acctatgtat	840
atttacatca gtcaccccaa accagaaggc ctggctgctg ccaagaagct ttgtgagaat	900
ctttgcaaa cagttcatgc tgaataactct agatttgcata atcagattaa tactgctgta	960
cctttaccag gctatacaca accctctgct ataagtagtg tccctctca accaccat	1020
tatccatcca atggctatca gtctggttac cctgttgttc cccctctca gcagccagtt	1080
caacccctt acggagtacc aagcatagtg ccaccagctg tttcattagc acctggagtc	1140
ttgccccat tacctactgg agtcccacct gtgccaacac aatacccgat aacacaagtg	1200
cagccctccag ctagcactgg acagagtccg atgggtggtc cttttattcc tgctgctct	1260
gtcaaaaactg ctttgccctgc tggccccag ccccgcccc agccccagcc cccactccca	1320
agtcagcccc aggcacagaa gagacgattc acagaggagc taccagatga acggaaatct	1380
ggactgcttg gataccaggt taaataaaat accctgtttt cctatcttca ctttattctt	1440
ctactatatt ctccctttaa aaaagataaa ttcacatcat tctccctgta ctaggatttc	1500
tgctttctgg aattcatttt ggttaggtt tttatcctat tcaacagact cttgaaagcc	1560
tctgagagtt ctactttct tatacatctc actcaaagct cttgatctac cagtatgtgg	1620
tttgtattta aaaccttggc tttcagtggt gctctctttt ttaccctcca cctaaaaaaag	1680
agagtgatat ctccctccag tctccccacc cctcaagact gctagaaaaag gagtgattct	1740
gtacatgtaa ttgtaaaagtt agccactaaa gttaaaaaga ttcttaattt gtagtttgg	1800
tgcaattttt taagaagttac ctttccattt tgccagaatc cttgaatcat tctttaaacc	1860
aaagcattttt ttatagttt ctagctaggt ttatagaaac tagtggagct atggcagtc	1920
agttaaaaac aggccataga tagcataatg aattataaca cccctgtcca agtcctatag	1980
agaaaaaaaaaa aaatccctac ttttgactac agttacacag cagatcccaa agagctttgt	2040

110

agtagttaa cgtactacaa cttatcagaa agatgaggca cttgacagtt acattaagga 2100
gctaaaagtca atacggcagt tgttagatttgc ctaatgccac tgtatTTTC tgctcatagc 2160
atggacccat tcatatgact aatttaggta caggcttctc cagtcagaat gagattgaag 2220
gtgcaggatc gaagccagca agttcctcag gcaaagagag agagagggac aggcagttga 2280
tgccTCCacc agcTTTCCA gtgactggaa taaaaacaga gtccgatgaa aggaatgggt 2340
ctgggacctt aacagggagc catggtgagt gtgatatagc tgggggaaca ggggagtggc 2400
taagacttgt ctaaagctat tagTTTCTC agccgggcgc agtggctcac gcctgtaatc 2460
ccagcacttt gggaggccga ggtgggcaga tcacctaagg tcaggagttc aagaccagct 2520
tggccaacat agtgaardcc catctctact aaaaatacaa aaactagcgg gcatggtgtt 2580
gggcgcctgt aattccagct actcaggggg ttgaggcagg agaatcgctt caacctggga 2640
ggcagaggtt gcagtgagcc aagatcagac cactGCCCTC cagcctggc aatagagcaa 2700
gactccatct cataaataaa taaatacata aataaagcta ttaattttct aacctgatgt 2760
tcattcaggt gttaatcca acctctataa tctgttggcc agtggaaata ctTTTGGCT 2820
gggcacggtg gctcacgcct gtaatcccag cactttggga ggccaagggtg ggcggataac 2880
ctgaggctcag gagtttgaga ccagcgtggc taacacggtg aaACCCGTC tctactaaaa 2940
atagaaaaat taagctggc atggtggtgc atgcctgtaa ttccagcggc ttggaaaggct 3000
gaggcaggag aatcacttga acttgggagg tggagggtgc agtgagccaa gatcacacca 3060
ctgcattcca gcctggcac tagagtgaga ctctgtctca aaaaaaaaaaaga aagagaaaaga 3120
gaaaatagtt tctaaaaat tgtatacaga caaccttttta ttccaacaa acgtgtgccg 3180
agagagagag agagaaaata gtttaaaaa aattgtatac agacaacott ttgtttccaa 3240
ccaacgtgta tctagaaaag agttagtcg cttatTTTAT acatagcatc agtgaatagt 3300
aatgagtggt aggtcatttc aaaatcctgt tgcctatatt atgtgaatac caggaggtca 3360
tctgatacgg acttaataaa gtttgattt gctttatatt gggagctgag ccacacctcc 3420
ccttataact ctattggta gtaatggta gtttggct gtttaggaaaa tggcccttt 3480
tagcattcca gaactctaaa tcctgttagag gtacatggga tattttatTC ttgcctgta 3540
ctcataaaaa tgaacagaag aaaatacgtt ttttctttt cttaacttct tttcttttaa 3600
ctctttaaaa ggtgaaatat cagccctcaa gagactca tgcataactt ctttttttc 3660
tttttttttc ttttttttgc gtttctttt tctttctctg ttttcttaca tggttctgg 3720
ggattcacat ttgctgatgc tggtgctgtt tttcgtgtga tcttcaacgt ttttgggtga 3780
ccattgaccc tgcgtacccaa aaatgggtgc caactaacca cttaaaatTA acatctttt 3840
tttaattaac gaatttatgg tattttttt tttcccttgg cggggatggg gttgggggttg 3900

111

tttttctct	attctagatt	atccagccaa	gaagatgaaa	actacagaga	agggatttgg	3960
cttggggct	tatgctgcag	attcatctga	tgaagaggag	gaacatggag	gtcataaaaa	4020
tgcaagtgt	tttccacagg	gctggagttt	gggataccaa	tatccttcat	cacaaccacg	4080
agctaaacaa	cagatgccat	tctggatggc	tccctaggaa	acagtggAAC	agagtttga	4140
ccctcagtga	ctcttccttag	caataatgca	tgcatttgat	ttaacaagac	tctggggcct	4200
gtgctggaa	ccatctggac	ctttgcagaa	gttagagatt	cagtcccccc	ctttcttaaa	4260
ggggttccctt	aacaaccaca	aaaatcccta	tttctgcagt	ggcatagaat	ctgttaaaat	4320
ttaatttagaa	tcacaaattt	atctcagaag	cttttaaca	gttggtaaaa	tgtgcttgc	4380
caacaaagca	tcctaacagg	gtcggtccca	tacacatttgc	acctggtcag	cctttccag	4440
gtgaatagcc	ccagttctga	cataaagaaa	gttttatttg	tatTTacta	ctgttggtc	4500
aattttgata	tataactggt	tacaaacaga	gccttactat	ttatttagtgg	ggaaatgatt	4560
ttaagaccgt	ccttttcagt	atTTattct	gacagatctg	catccctgtt	ttgtttggaa	4620
ttatTTctgt	tttggaaaat	gtgtctcat	ttaaaaactgt	tggatatagc	tggatcctgg	4680
ataggaaaat	gaaatttattt	tttcattgtg	tttttaattt	ggggtgatcc	aaagctggca	4740
ccttcaggca	cattggtctc	atagccatta	ctgttttat	tgccttctta	agatcctgtc	4800
ttcagctggg	tcagagaaaa	cttcttgact	aaaactggtc	agaactcatc	acagaaatga	4860
aatacagtgg	tctctctctc	ccagaactgg	ttgcagctaa	aacagagaga	tctgactgt	4920
ggctatagga	tttggactt	aatgactgaa	attgcaaattt	gtccttttc	ttggcattac	4980
agatTTGCC	aaaataactt	tttgtatcaa	atattgtatgt	gtgaaagtga	aggagctgt	5040
ctgctgaacc	aggaatagtt	tgagatatttgc	aactgtcatt	tttgcacatt	tgaataacttt	5100
gcaggctggc	tttgtataaa	cttacctct	ggtttcttat	atgttgtaaa	tatTTagacc	5160
ataatttcat	tataaataaa	tctataaata	ttc			5193

<210> 84
<211> 5410
<212> DNA
<213> Homo sapien

<400> 84	cacccagctt	gccccattga	tttggataga	cgagaccaca	tttggttata	acctattaga	60
	caaccagcaa	aaaggaatttgc	tgttcttaaa	tattgggtta	tctcaataga	ctggaaaata	120
	atggtagtga	atccataaga	acagatttgc	tctgaactct	gtggggcagt	tgacttttaa	180
	aatcaagaag	tttatttgc	ccccagctgt	taaagcatgt	tctgagacat	tagctagcaa	240
	aaacccttct	aatgtatata	gtcccttttgc	gttatttggaa	aacttgaaac	attacaaattt	300

agatctggtg acccatttga ctcctttttt gggtttgacc acattgagat catccttttt	360
tgtttcttctt tatagggatc gtccattata tcttcatgtt cagggccaga cacggaaatt	420
agtggacagt aagtaatgtt ttggctcac gtagcacttt ttgtgaagag caaagtacag	480
ggctctgtat agcaagttgg caaagtgtcc ctgatggctg ttctaaccct tgttcatgaa	540
ctatacacga atttgtatgg gagtttagag ggatggagag ccacatattt gggtaacgta	600
taaaggcagat ttacggtgaa taattgaaca ctggcctgcc tgggcactag tcagttcatt	660
ccattcagtt ttactgtctg tgttttctta aggagctgta aaccggatca aagaaattat	720
caccaatgga gtggtaaaag ctgccacagg aacaagtcca acttttaatg gtgcaacagt	780
aactgtctat caccagccag caccatcg tcagttgtct ccagctgtta gccagaagcc	840
tcccttccag tcagggatgc attatgttca agataaaatta tttgtgggtc tagaacatgc	900
tgtacccact ttaatgtca aggagaaggt ggaaggcaca ggctgctcct atttgcagca	960
cattcagatt gaaacaggtg ccaaagtctt cctgcggggc aaaggttcag gctgcattga	1020
gccagcatct ggccgagaag ctttgaacc tatgtatatt tacatcagtc accccaaacc	1080
agaaggcctg gctgctgcca agaagctttg tgagaatctt ttgcaaacag ttcatgctga	1140
atactctaga tttgtgaatc agattaatac tgctgtacct ttaccaggct atacacaacc	1200
ctctgctata agtagtgtcc ctccctcaacc accatattat ccatccaatg gctatcagtc	1260
tggttaccct gttgttcccc ctccctcagca gccagttcaa cctccctacg gагtaccaag	1320
catagtgcca ccagctgttt cattagcacc tggagtcttgc ccggcattac ctactggagt	1380
cccacctgtg ccaacacaat acccgataac acaagtgcag cctccageta gcactggaca	1440
gagtccgatg ggtggtcctt ttattctgc tgctcctgtc aaaactgcot tgctgctgg	1500
ccccccagccc cagcccccagc cccagccccc actcccaagt cagccccagg cacagaagag	1560
acgattcaca gaggagctac cagatgaacg ggaatctgga ctgcttggat accaggttaa	1620
ataaaaatacc ctgttttctt atcttcaact tattttctta ctatattctc cctttaaaaa	1680
agataaaattc acatcattct cccagacta ggatttctgc tttctggaat tcatttttgtt	1740
tagttttttt atccattca acagacttt gaaaggctct gagagttctt actttcttat	1800
acatctcaact caaagctctt gatctaccag tatgtggttt gtattttaaa ccttggcttt	1860
cagtggtgct ctctctttta ccctccacct aaaaaagaga gtgatatctc cctccagtc	1920
ccccacccct caagactgct agaaaaggag tgattctgta catgttaattt gaaagttagc	1980
cactaaagtt aaaaagattc ttaattttgtt gttttgggtc aattttatca gaagtacctt	2040
tccatTTTgc cagaatcctt gaatcattct taaaacccaaa gcattttttt atagttctta	2100

113

gctaggttta tagaaactag tggagctatg ggcagtcgt taaaaacagg ccatagatag	2160
cataatgaat tataacaccc ctgtccaaagt cctatagaga aaaaaaaaaa tccctacttt	2220
tgactacagt tacacagcag atcccaaaga gctttagt agtttaacgt actacaactt	2280
atcagaaaaga tgagggactt gacagttaca ttaaggagct aaagtcaata cggcagttgt	2340
agatttgcta atgccactgt attttctgc tcatacgatg gaccattca tatgactaat	2400
ttaggtacag gottctccag tcagaatgag attgaaggtg caggatcgaa gccagcaagt	2460
tcctcaggca aagagagaga gagggacagg cagttgatgc ctccaccagg ctttccagtg	2520
actggaataa aaacagagtc cgatgaaagg aatgggtctg ggaccttaac agggagccat	2580
ggtgagtgta atatacgctgg gggAACAGGG gagttggctaa gactggctta aagctattag	2640
ttttctcagc cgggcgcagt ggctcacgccc tgtaatcccc gcactttggg aggccgaggt	2700
gggcagatca cctaaggcata ggagttcaag accagcttgg ccaacatagt gaaatccccat	2760
ctctactaaa aatacaaaaaa ctagcgggca tgggtgggg cgcctgtat tccagctact	2820
cagggggttt aggcaggaga atcgcttcaa cctgggaggc agaggttgca gtgagccaag	2880
atcagaccac tgccctccag cctgggcaat agagcaagac tccatctcat aaataaataa	2940
atacataaaat aaagctatta attttctaac ctgatgttca ttcaagggttt taatccaacc	3000
tctataatct gttggccagt gaaaatactt ttgggtggg cacgggtggct cacgcctgtat	3060
atccccagcac tttgggaggc caaggtgggc ggataacctg aggtcaggag tttgagacca	3120
gctggctaa cacggtaaa ccccgctct actaaaaata gaaaaattaa gctgggcatg	3180
gtgggtgcatt cctgtatcc cagcggcttg gaaggctgag gcaggagaat cacttgaact	3240
tgggagggtgg aggttgcagt gagccaagat cacaccactg cattccagcc tgggacttag	3300
agtgagactc tgtctcaaaa aaaaagaaag agaaagagaa aatagtttct aaaaattgt	3360
atacagacaa ccttttattt ccaacaaaacg tgtgccgaga gagagagaga gaaaatagtt	3420
ttaaaaaaaat tgtatatacaga caacctttt tttccaaacca acgtgtatct agaaaagagt	3480
tagtcgactt attttataca tagcatcagt gaatagtaat gagttggtagg tcatttcaaa	3540
atccctgtgc ctatattatg tgaataccag gaggtcatct gatacggact taataaaggt	3600
tgatTTTGTCT ttatattggg agctgagcca cacccccc tataactcta ttgggtcagta	3660
atggtcagtt tgtggctgtt aggaaaatgt tgccttttag cattccagaa ctctaaatcc	3720
tgttagaggtt catgggatat ttattcttt gcctgtactc ataaaaatga acagaagaaa	3780
atacgttttt ttcttttctt aacttctttt cttaactc ttAAAAGGT gaaatatcag	3840
ccctcaagag actcacttgc taactttctt ttttttcttt ttttttcttt ttgggtgttt	3900
tctttttctt ttctctgttt ttttacatgg ttctgggaa ttcacatttg ctgatgtgg	3960

114

tgctttttt cgtgtgatct tcaacgtttt tggtgacca ttgaccctgt gacctaaaaa 4020
 tggtgtccaa ctaaccactt aaaattaaca tctttttttt aattaacgaa tttatggtat 4080
 tttttttttt cccttggcg 99atggggtt ggggtgtt tttctctatt ctagattatc 4140
 cagccaagaa gatgaaaact acagagaagg gattggctt ggtggcttat gctgcagatt 4200
 catctgatga agaggaggaa catggaggtc ataaaaatgc aagtagttt ccacagggct 4260
 ggagtttggg ataccaatat ctttcatcac aaccacgagc taaacaacag atgccattct 4320
 ggatggctcc ctaggaaaca gtggAACAGA gtttgaccc tcagtgactc ttcttagcaa 4380
 taatgcatgc atttgattta acaagactct gggcoctgtg ctgggaacca tctggacctt 4440
 tgcagaagtt agagattcag tgccccccctt tcttaaaggg gttccttaac aaccacaaaa 4500
 atccttattt ctgcagtggc atagaatctg taaaattta attagaatca caaatttatac 4560
 tcagaagctt ttaaacagtt ggtgaaatgt gcttgcCAA caaagcatcc taacagggtc 4620
 gttcccatac acatttgacc tggtcagcct tttccaggtg aatagccccca gttctgacat 4680
 aaagaaaagtt ttatttgtat ttactactg tttggtcaat tttgatataat aactggttac 4740
 aaacagagcc ttactattta ttagtgggaa aatgattttt agaccgtcot tttcagttatt 4800
 taattctgac agatctgcat ccctgtttt gtttggatta tttctgtttt ggaaaatgtct 4860
 gtctcattta aaactgttgg atatacgctgg atcctggata ggaaaatgaa attattttt 4920
 cattgtttt ttaatttggg gtgatccaa gctggcacct tcagggcacat tggtctcata 4980
 gccattactg ttttattgc ctttctaaga tcctgtcttc agctgggtca gagaaaactt 5040
 cttgactaaa actggtcaga actcatcaca gaaatgaaat acagtggctc ctctctccca 5100
 gaactggttg cagctaaaac agagagatct gactgctggc tataggattt tggacttaat 5160
 gactgaaattt gcaaattgtc cttttcttg gcattacaga ttttgcAAA ataacttttt 5220
 gtatcaaata ttgatgtgtg aaagtgaagg agcttagtctg ctgaaccagg aatagtttga 5280
 gatattgaac tgtcattttt gcacatttga atactttgca ggctggcttt gtataaaactt 5340
 atcctctgggt ttcctatatg ttgtaaatat ttagaccata atttcattat aaataaaatct 5400
 ataaatattc 5410

<210> 85
 <211> 5271
 <212> DNA
 <213> Homo sapien

<400> 85
 ggcaaaaaaa aatttaattt tggcagattt tgctatttag aatctttgaa gttttctttt 60
 gtttaagatgg attgcattttt acttttgact aaaattcca gaattatgtg tggactctt 120

atatactggta tggtaaggc ctaactccctt acaataaaaa tttaaaatta agtaacaaaa	180
tctagataat ctttatttcc taaccatcat ctgtttggc actctaccac ctaggtacca	240
taactgaaca acctatatgc tctggtttc taacttttt actagttgtg ctaatattt	300
tacttggtag tcaaaggaaa tgaagtata atagtttcc ttctcttaca gagactttag	360
aagacagctt tccatcgaca ctaggccctt tagaccagcc tctgaaggaa atcctagcga	420
tgatcctgag ctttgccag cacatcgca ggacttgga gagaggctt atcctcggt	480
acaagaatg caaccagtgt gtactgtgt a ctttttttcc tcttaactgg ctgtgttatt	540
tttttgcctt tgaattaaag atgggttttta aaaaaatttgc tagctactg gcaaaaccac	600
atgattgagg actcttttta gctctgcagt aagaaggaag ctcaggagag aataaggcag	660
tgtttactga aggtaactg tatacttgc atttattttt tcctccaccc accttatgtt	720
gggacactgt ctccattcta ctttccttgc atgctaaaga atttgctgtg cattatattt	780
attgtatctt catttgaaca aattattact acttttggag cagactttat ctgttagcaa	840
gctgttagttg gaacacatta atgctgatta gtattgcagg aagaatttttta ttttgaatgt	900
tctatcaaga gttttctttt atatgatatg aaacacaaat tagttatgtt ttttgcctt	960
atgcataact gtatgtacac attttatgtt gaaagaataa cagaaaaacta ttatttcttc	1020
cagcaagtcc tcagccttaa acataggat atttttctct accctacccc cttcttttt	1080
tctaccctaa ataaaagata ttctggctc tctgtatgc aaaaaaaaaatggaaatttgc	1140
tatatgtatg tttaactcag agatataaaaa aaacctaaaa agaaaaacttg tcatacaaat	1200
attataagta gccttaacaa gatgtggatc tgcatggact gtttattccc tgccaagttt	1260
ctctataattt gatcttccag tttcataaaaa gaccttactg gttctgaaat tttgtatgg	1320
ttacccaagt ttcttattttt attttttttt taaataaaaatggatgtt taatttagaca	1380
agaggtttta gagagtagtc aagtaacatt tggtcatcat ttacaggcat ttgcaagttt	1440
aatcaactggc atgttggatc aattatcccc agctcagctg cttctcccttc tagcaagtga	1500
ggattctctg agagcaagag tggatgaggc catggaaactc attattgcac atggacggga	1560
aaatggagct gatagtatcc tggatcttgg atttagtagac tcctcagaaa aggtacagca	1620
ggaaaaaccga aagcgccatg gctctagtcg aagtgttagta gatatggatt tagatgatac	1680
agatgtatgtt gatgacaatg cccctttttt ttaccaacctt gggaaaaagag gattttatac	1740
tccaaaggcct ggcaagaaca cagaagcaag gttgaattgt ttcagaaaca ttggcaggat	1800
tcttggacta tgtctgttac agaatgaact atgtctatc acattgaata gacatgtat	1860
taaagtatttgc ttggtagaa aagtcaattt gcatgatggatc gcttttttttgc atcctgtat	1920

116

gtatgagagt ttgcggcaac taatcctcg	gtctcagagt tcagatgctg atgctgttt	1980
ctcagcaatg gatttggcat ttgcaattga cctgtgtaaa gaagaaggtg gaggacaggt		2040
tgaactcatt cctaataatggtg taaatataacc agtcactcca cagaatgtat atgagtatgt		2100
gcggaaatac gcagaacaca gaatgttggt agttgcagaa cagcccttac atgcaatgag		2160
gaaaggcgtcta cttagatgtgc ttccaaaaaaaa ttcatggaa gatttaacgg cagaagattt		2220
taggcctttt gtaaatggct gcggtaagt caatgtgcaa atgctgatca gtttacctc		2280
tttcaatgtat gaatcaggag aaaatgtga gaagcttctg cagttcaagc gttggttctg		2340
gtcaatagta gagaagatga gcatgacaga acgacaagat cttgtttact tttggacatc		2400
aagcccatca ctgccagcca gtgaagaagg attccagcct atgcctcaa tcacaataag		2460
accaccagat gaccaacatc ttccctactgc aaatacttgc atttctcgac tttacgtccc		2520
actctattcc tctaaacaga ttctcaaaca gaaattgtta ctcgcccatta agaccaagaa		2580
ttttggtttt gtgttagagta taaaaagtgt gtattgctgt gtaatattac tagcaaattt		2640
tgtagatttt ttccatttt tctataaaag tttatggaaag ttaatgctgt catacccccc		2700
tggtggtacc ttaaagagat aaaatgcaga cattccttgc tgagttataa gcttaaaggc		2760
ctaaggagca cttagcaacat ttggctataat tggggcata gtcaccaact tctgggtcta		2820
accccaagcca aagatgacag cagaacaaca taatttacac tgtgatttat cttttgctg		2880
agggggaaaa aatgtaaatg ttctgaaaat tcactgctgc ctttggaaactgtttcag		2940
caaagggtct tgtatagagg gaataggaa ttcaaaata aaaaattaag tatgttctgt		3000
gttttcattt taactttttt tatgggtttt aatttgggtt tggctgcaac tgtgtatcat		3060
gtatatggaa cttgtaaaaa agttctcgac attcagatct taagagatga aatcactttt		3120
acctataaaa accacttttta ttgcggtttg actgcattga gctctaggat attaaatgat		3180
atcactaata ttttgcattgt aatttgcata tttgagtgag ggcactttt ttgtacat		3240
gatggggcca atgcacaata cttttatcac aatcaacttt ttctttgtat ccctatttca		3300
atgagcagtc agtctcaaga ggttactgca ctgcattct aactagacat ttgtactaag		3360
gtatttcagt tatgtaaaact cagcctgggc actttctgt aactgtaaaa tgtttataaa		3420
gatcatgatt attgaagata catggaa aattttaat gttcgtgagc agcttaacta		3480
cttttgtatc tagcctttt taagtatctt gttacattta cttttttaaa taaagaaattt		3540
acagaagaaa tgtcaagtaa tattgaagaa acaatagttt ttatgtatgt agttgtacat		3600
ttttaaacta agggcaatac actgacatgg ttatgtgcat aaaaattttg acttaaagaa		3660
ctggaaagttt atatacacct ggactataag aaacagaaga aaatcagtcc acattttaca		3720
gttagcagag aatcctaaat ggcactggcc tggccacctt ttcattttac aaatggggga		3780

agtgaagtgt gacccttac ttggcatagg aagttaactt acacctaata actgacaggt	3840
ttttgtttt atgacctatt aattatgtag cctaggatta atatccaaa attactctgg	3900
tttaagtagc tttattcagt ggcataataa cactgtttc ttcccttaagt cttaatgaa	3960
gtgacttaaa acagtcactt tacatattaa aaatgaggag agcaattctc tggaatctct	4020
ccttcagtt cctttagg atttctggcc ttgaggatag tcttcatgtt caaaggcact	4080
atgcttttat tatataactt cttcagaag actgaaccac atgatattct cagccctgtt	4140
aacactaaaa atattnaaaa ctgaatgata gtatgtactc attgtattac ttaaaactta	4200
tataacacgc tgtatttagat gtgtgtaaat tagccaaagg ttattnaca aagtgagaca	4260
ttggtttta tgtctaaatg ctatttctga ataaatgaaa tagtaattag atcaagagct	4320
gattagcatc aatgtgtttt aaagatataa aatttataca tcaccttaac ctctgtatgc	4380
acatgatggg attgataaaa tattaaatga gaacaaacta gatatgatta ggacatttg	4440
aaccctaatt gtgaatttat tttaatagt tactgaaatg aaaatattta aaataatgca	4500
caatgtctta agtcttccta aatcaagatt ttggtaaaaa aatacttcta ataatagtaa	4560
aagattnnnn tttaagtaa atcataaaac ggttctaaat gtaaaataaa gacatgtaaa	4620
ataaagttct ctttggct ttttagtgt tttaatctaa caattgaaaa caaattttagg	4680
aagagaagac caagaatgaa cttaactgag tttttcaga gtttgctact actattnnn	4740
tccctaaatc atctggatac caagactatc cagtaaaatg gataactggg gcagacttga	4800
gagggtatnt taaaggaatg atttcaactat ttagtagctg cccccaaaca acatccctcc	4860
cataaaagata ctattnnac attttaaagg tagtcagcaa ttccatgtt taaactcaag	4920
ttgagataat cccttgaggc agtagttcc atgcttctgt atgttgaag attcatttgt	4980
aaagtttgtt aatgcagatt cttaagcatt cctcatcctc ttgcctcctt tctgattcag	5040
taagtctttg gtggaggcca ggaatctca tgcagatcat cccaggtgat tctgaaacac	5100
tgcccaaaga atattnctt ttatttaca aatataaatg tcccgtgaa agctcctgag	5160
agccaaacct ttccactta gaactgctta caatctatgg aaaagtacat ctattgataa	5220
actagtccta gtttggattc ttccactga taaggggctg gttgaaatg c	5271

<210> 86
<211> 3159
<212> DNA
<213> Homo sapien

<400> 86	
tgggttgacc gatgctggc agctgagcgg accaatcgcc cccctagact gagacgttgg	60
cgttggaaat cagccaatgg caggtctaca ctggagctc ctctccgcct cttcgccct	120

gcctgcgagt gttctgaggg aagcaaggag gcggcggcgg ccgcagcgag tggcgagtag	180
tggaaacgtt gcttctgagg ggagcccaag gttagggaggc gaggcgacgg tgtgcgggag	240
cgggctctcc agggacttcc cgggtccgca actggcaggg ccgttcgatt cgcaaaaaat	300
cccgttcgt ttctgttgtt ttccctttat ttttaggagt gcccggggcg acgggacccc	360
gggagagggg aaagggaaaca gtctggggtc cgggcatacg tttggccgg gctgggtta	420
gggggacggc ggtgcgggct gggccgggtt gggcgccgg gggggccggat gatggggcga	480
gtccggacct tggcgggcga gtgctcgccg caggcgcaag cgcaagatct ctcgcggtc	540
gtcctctcgg cccctccctc tggggggacc cccagtgcctt ggctgtcagt ggcagcccc	600
agcccgccgg accccctgggg actctggcgc cctgttctgc agatgaccgg ttctaaccgg	660
ttcaagctga accagccacc cgaggatggc atctcctccg tgaagttcag ccccaacacc	720
tcccagttcc tgcttgtctc ctccctggac acgtccgtgc gtctctacga tttggccggcc	780
aactccatgc ggctcaagta ccagcacacc ggccgggtcc tggactgcgc cttctacgt	840
ccaacgcatg cctggagtgg aggactagat catcaattga aaatgcataa tttgaacact	900
gatcaagaaa atcttgggg gacccatgtt gcccctatca gatgtgttga atactgtcca	960
gaagtgaatg tggatggtcac tggaaagtgg gatcagacag taaaactgtg ggatcccaga	1020
actccttgta atgctgggac cttctctcag cctgaaaagg tatataccct ctcagtgtct	1080
ggagaccggc tgattgtgg aacagcaggc cgcaagatgt tgggtgtggg cttacggAAC	1140
atgggttacg tgcagcagcg cagggagttcc agcctgaaat accagactcg ctgcatacg	1200
gcgtttccaa acaaggcaggg ttatgtatta agctctattt aaggccgagt ggcagtttag	1260
tatgggacc caagccctga ggtacagaag aagaagtatg ctttcaaattt tcacagacta	1320
aaagaaaaata atattgagca gatttaccca gtcaatgcc tttctttca caatatccac	1380
aatacatttgc ccacaggtgg ttctgtatggc tttgtaaata tttggatcc atttaacaaa	1440
aagcgactgt gccaaatttcca tcgggtacccc acgagcatcg catcaatttc ctccatgt	1500
gatgggacta cgcttgcattt agcgtcatca tataatgtatg aaatggatga cacagaacat	1560
cctgaagatg gtatcttcat tcggcaagtg acagatgcag aaacaaaacc caagtcacca	1620
tgtacttgac aagatttcat ttacttaatg gcatgttga tgataataaa acaattcgta	1680
ctcccaatg gtggatttat tactattaaa gaaaccaggg aaaatattaa ttttaatatt	1740
ataacaaccc gaaaataatg gaaaagaggt ttttgaattt tttttttttaa ataaacaccc	1800
tcttaagtgc atgagatggt ttgatggttt gctgcattaa aggtatgg gcaaaacaaa	1860
ttggaggggca agtgcactgca gttttgagaa tcagtttga ctttgatgtat tttttgtttc	1920

119

cactgtggaa ataaatgttt gtaaataagt gtaataaaaa tcccttgca ttctttctgg	1980
accttaaatg gtagaggaaa aggctcgta gccatttgg tctttgctg gttatagttg	2040
ctaattctaa agctgcttca gactgcttca tgaggaggtt aatctacaat taaacaatat	2100
ttcctcttgg ccgtccattt tttctgaag cagatggttc atcatttcct gggctgttaa	2160
acaaagcgag gttaaggta gactcttggg aatcagctag tttcaatct tattagggtg	2220
cagaaggaaa actaataaga aaacccctta atatcattt gtgactgtaa acaattat	2280
attagcaaac aattgatccc agaaggcaaa attgttttag tcaagtaatga gctgagaaaa	2340
gacagagcat atctgtgtat ttggaaaaat aattgtaacg taattgcagt gcatttagac	2400
aggcatctat ttggaccctgt ttctatctct aatgaattt ttggaaacat taatgaggtt	2460
tacatatttc tctgacatatt atatagttct tatgtccatt tcagttgacc agccgctggt	2520
gattaaagtt aaaaagaaaa aaattatagt gagaatgaga ttcatctaa tgtaatgcac	2580
taaagcagaa cacgaactta gcttggccta ttctaggtag ttccaaatag tattttgtt	2640
gtcaaacttt aaaatttata ttaatttgc aatgtatgtc tctgagtagg acttggac	2700
ttccctgagat ttatttatc cgtgatgtat ttttttaat tctttgata cagagaaggg	2760
tctttttttt tttaagtatt tcagtgaaaa ctgggtgtaa gtctgaaccc atctttgaa	2820
atgtattttc ttcattgcag gtccaccta tcattcctgt aaagtggttt ctctatggaa	2880
agctttgttt gtttcctaca aatacatgtc tattcctaa gggatgtgtt agagttactg	2940
tggatttctc tgccccctgt cttacaagaa acttgcattt gtaccttaat actttgttt	3000
ggatgaggag tctttgtgtc cctgtacagt agtctgacgt atttccctt ctgtcccccta	3060
gtaagccccag ttgctgtatc tgaacagttt gagctttt tgtaatatac tctaaacctg	3120
ttatccctgt gctaataaac gagatgcaga acccttgaa	3159

<210> 87
<211> 1018
<212> DNA
<213> Homo sapien

<400> 87	
gcccttagcg tggtcgcggc cgaggtaacc tgccccgttc ttatgtctcg aatgtccaa	60
cctgaagctg aagaagccgc cctgggtgca catgccgtcg gccatgactg tgtatgtct	120
ggtgtgggtg tcttacttcc tcattcacccg aggaataatt tatgtgtta ttgttgaacc	180
tccaagtgtc ggttctatga ctgtatgttca tggccatcag aggccagtag ctttcttggc	240
ctacagagta aatggacaat atattatggaa aggacttgca tccagcttcc tatttacaat	300
gggaggttta ggtttcataa tcctggaccg atcgtatgtca ccaaataatcc caaaactcaa	360

120

tagattcctt	cttctgttca	ttggatcg	ctgtgtccta	ttgagtttt	tcatggctag	420
agtattcatg	agaatgaaac	tgccgggcta	tctgatgggt	tagagtgcct	ttgagaagaa	480
atcagtggat	actggat	tttgc	tcctgtcaa	tgaagttta	aaggctgtac	540
atatgaaatg	tggaaaagaa	tgaagagcag	cagtaaaaga	aatatctgt	aaaaaaacag	600
gaagcgtatt	gaagcttgg	ctagaattt	ttcttggtat	taaagagaca	agtttatcac	660
agaatttttt	ttcctgctgg	cctattgcta	taccaatgat	gttgagtg	tttttttt	720
tagttttca	ttaaaatata	ttccatata	acaactataa	tatcaaataa	agtgattatt	780
ttttacaacc	ctcttaacat	ttttggaga	tgacatttct	gat	ttttcaga	840
aaatccagaa	gcaagattcc	gtaagctgag	aactctggac	agttgatcag	cttacctat	900
ggtgctttgc	ctttaacttag	agtgtgtat	ggtagattat	ttcagatatg	tatgtaaaac	960
tgtttcctga	acaataagat	gtatgaacgg	agcagaaata	aatactttt	ctaattaa	1018

<210> 88
<211> 2075
<212> DNA
<213> Homo sapien

<400> 88						
ggcggttccg	tacagggtat	aaaagctgtc	cgcgccggag	cccaggccag	ctttggggtt	60
gtccctggac	ttgtcttgg	tccagaacct	gacgacc	cgacggc	gtcttttg	120
actaaaagac	agtgtccagt	gctccagc	aggagtctac	ggggaccg	tccgcgc	180
ccaccatgcc	caacttctct	ggcaactgga	aaatcatccg	atcgaaaac	ttcgaggaat	240
tgctcaaagt	gctgggggt	aatgtgtatgc	tgaggaagat	tgctgtgg	gcagcgtcca	300
agccagcagt	ggagatcaa	caggagggag	acacttcta	catcaaaacc	tccaccaccg	360
tgcgcaccac	agagattaac	ttcaagg	gggaggagtt	tgaggagc	actgtggatg	420
ggaggccctg	taagagcctg	gtgaaatgg	agagtgagaa	taaaatgg	tgtgagcaga	480
agctcctgaa	gggagagggc	cccaagac	cgtggacc	agaactgacc	aacgatgggg	540
aactgatcct	ggtaagtcc	gcctcctccc	cactaatagc	aaacccag	ctac	600
agattctctg	ggagacccc	gggtgcag	gactcaagaa	caaccatgg	tggactcc	660
accctgctga	tgggactg	tgaacagaac	taaggtgt	ctatcccata	cagtgc	720
tgtgaattag	aaatgggtt	cctttatgc	aagcaaagg	catgtactg	gggatccc	780
cagttctca	gggagatctt	cctggcttga	ggaggaggac	ggccccagg	ggctctatt	840
ctatcctccc	tccattgat	cctggcatt	ctgggacc	ctcctgc	ttggtcttga	900
gccaagaagc	agg	tttggac	ctggaggc	agcagagtac	ctccattaa	960

121

caaagccaca ggaccccagg ggcctctcag gctaacaact acttctgtcc ttccagacca	1020
tgacggcgga tgacgttgtg tgcaccaggg tctacgtccg agagtgagtg gccacaggtt	1080
gaaccggcgc cgaagcccac cactggccat gtcacccgccc ctgcttactt gccccctccg	1140
tcccacccccc tccttcttagg atagcgctcc ctttacccca gtcacttctg ggggtcactg	1200
ggatgcctct tgcaagggtct tgctttctt gacctcttct ctccctccctt acaccaacaa	1260
agaggaatgg ctgcaagagc ccagatcacc cattccgggt tcactccccg cctccccaag	1320
tcagcagtcc tagccccaaa ccagccaga gcagggtctc tctaaagggg acttgagggc	1380
ctgagcagga aagactggcc ctctagcttc taccctttgt ccctgttagcc tatacagttt	1440
agaatattta ttgttaatt ttattaaaat gcttaaaaaa aataaaaaaaa aaaaaacaaa	1500
aaaaaaaaaaag aagagcccg cgcgcaaaac ccgcgtggcc atggcgccgc gaccgcggg	1560
gcgcgaaaaac agtggcgtac ctgcggcct cccaaattt tccccacca cctttagcgc	1620
agcgaccaac gtgcgcgcgg cgcaagcgaaa gcggccgcga cgagcgccgg acgctacgcg	1680
acggacggcg cgggcccggca ccacgcacc accgtcacggg cagccgcag cgacgcggc	1740
ggcggcgcct gtcacaacc gaggtctgcc tagttgctgc tcccggtgcc gagccaaggc	1800
ccgctacgca cggccacgca gggctgaggc agcggcacgc gcgcggcgtg caacgcggc	1860
ggcacccggc tggagggggg gaggcacccgc aacacggccg acgcggcgaa gagcgggaaac	1920
aaacgcacac gacccacacc gcaacgggtga gcaacgaccg agcggccagc ggccgaccgc	1980
gcgtggcagc aggccgacgc gcaacgagac gcgcgagagc gagagaccac tccgaggcgc	2040
cggcccggt gtgccaggcc cgacgcgtgg tggcc	2075

<210> 89
<211> 1557
<212> DNA
<213> Homo sapien

<400> 89	
gccccacccca agccggtttc acaaactccg tttcttacccg taaggtttct cccctctcgc	60
cgctcgggca agctgatcac aggtgtgtcg ggagcctagg agtctacggg gaccgcctcc	120
cgcgccgcca ccatgccccaa cttctctggc aactggaaaa tcatccgatc ggaaaacttc	180
gaggaattgc tcaaagtgtct ggggtgaat gtatgtgtga ggaagattgc tgtggctgca	240
gcgtccaaac cagcgtggaa gatcaaaccg gagggagaca ctttctacat caaaacctcc	300
accaccgtgc gcaccacaga gattaacttc aagggtgggg aggagtttga ggagcagact	360
gtggatggga ggcctgttaa gagcctgttg aaatgggaga gtgagaataa aatggtctgt	420
gagcagaagc tcctgaaggg agagggcccc aagacctcgt ggaccagaga actgaccaac	480

122

gatggggAAC	tgatcctgac	catgacggcg	gatgacgttg	tgtgcaccag	ggtctacgtc	540
cgagagttag	tggccacagg	tagaaccgcg	gccgaagccc	accactggcc	atgctcacgg	600
ccctgcttca	ctgccccctc	cgtcccaccc	cctccttcta	ggatagcgct	ccccttaccc	660
cagtcacttc	tgggggtcac	tgggatgcct	cttgcagggt	cttgctttct	ttgacctctt	720
ctctcctccc	ctacaccaac	aaagaggaat	ggctgcaaga	gcccagatca	cccattccgg	780
gttcaactccc	cgcctcccca	agtcagcagt	cctagccca	aaccagccca	gagcagggtc	840
tctctaaagg	ggacttgagg	gcctgagcag	gaaagactgg	ccctctagct	tctacccttt	900
gtccctgttag	cctatacagt	ttagaatatt	tatttgttaa	ttttattaaa	atgctttaaa	960
aaaataaaaaa	aaaaaaaaaca	aaaaaaaaaa	agaagagccc	ggcgcgcgaa	acccgcgtgg	1020
ccatggcgcg	gcgaccccgcg	gggcgcgaaa	acagtggcgt	acctcgccgc	ctccccaaat	1080
tctccccacc	caccttttagc	gcagcgacca	acgtgcgcgc	cgcgcagcgg	gggcggccgc	1140
gacgagcgcc	ggacgctacg	cgacggacgg	cgcggccgg	caccacgcca	ccacgtcacf	1200
ggcagccgccc	agcgcacgccc	cggcggccgc	ctgctcacaa	ccgaggtctg	cctagttgt	1260
gtcccccgtg	ccgagccaag	gcccgcctacg	cacgcacacg	cagggctgag	gcagcggcac	1320
gcgcgcggcg	tgcaacgcgg	gcccgcacccg	gctggaggggg	gggaggcacc	gcaacacggc	1380
cgacgcggcg	aagagcggga	acaaacgcac	acgacccaca	ccgcaacggt	gagcaacgac	1440
cgagcggccca	gcggcgacccg	cggcgtggca	gcaggcgacg	acgcccacgag	acgcgcgaga	1500
gcgagagacc	actccgaggc	gccggcccccgg	gtgtgccagg	cccgacgcgt	ggtgccc	1557

<210> 90
<211> 1430
<212> DNA
<213> Homo sapien

<400> 90	ggcggttccg	tacagggtat	aaaagctgtc	cgcgcggag	cccaggccag	ctttggggtt	60
	gtccctggac	ttgtcttgg	tccagaacct	gacgacccgg	cgacggcgcac	gtctcttttg	120
	actaaaagac	agtgtccagt	gctccagcct	aggagtctac	ggggaccgccc	tcccgcgccg	180
	ccaccatgcc	caacttctct	ggcaactgg	aaatcatccg	atcgaaaac	ttcgaggaat	240
	tgctcaaagt	gctgggggtg	aatgtgatgc	tgaggaagat	tgctgtggct	gcagcgtcca	300
	agccagcagt	ggagatcaaa	caggagggag	acactttcta	catcaaacc	tccaccacccg	360
	tgcgccaccac	agagattaac	ttcaagggtt	gggaggagtt	tgaggagcag	actgtggatg	420
	ggaggccctg	taagagcctg	gtgaaatggg	agagtgagaa	taaaatggtc	tgtgagcaga	480
	agctcctgaa	gggagagggc	cccaagaccc	ctaggatagc	gctccccctt	ccccagtcac	540

123

ttctgggggt cactgggatg cctcttgcag ggtcttgctt tctttacccct cttcttcct
ccccctacacc aacaaaagagg aatggctgca agagcccaga tcacccattc cgggttca
ccccgcctcc ccaagtcagc agtccttagcc ccaaaccagc ccagagcagg gtctctctaa
aggggacttg agggcctgag cagggaaagac tggccctcta gcttctaccc tttgtccctg
tagcctatac agtttagaat atttatttgt taattttatt aaaaatgcttt aaaaaaataa
aaaaaaaaaaa acaaaaaaaaaaaa aaaagaagag cccggcgcgca gaaaccccgcg tggccatggc
gcggcgacccc gcggggcgcg aaaaacagtgg cgtacccgcgc ggctccccca aattctcccc
acccacccccc agcgcagcga ccaacgtgcg cgccgcgcag cgggggcggc cgcgacgagc
gccggacgct acgcgcacgga cggcgccggc cgccaccacg ccaccacgta acgggcagcc
gccagcgcac gccccggcg aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa
gtgccgagcc aaggccccgt acgcacgccc acgcagggct gaggcagcgg cacgcgcgcg
gcgtgcaacg ccggcgacac ccggctggag gggggggaggc accgcaacac ggccgcacgcg
gcgaagagcg ggaacaaaacg cacacgaccc acacccgcaac ggtgagcaac gaccgagcgg
ccagcggcgca ccgcggcggtg gcagcaggcg acgacgcccac gagacgcgcg agagcgcag
accactccqaa qqcgcggggcc cqqaqtgtgcc aqgccccqacq cgtgggtggcc
1430

```
<210> 91  
<211> 1265  
<212> DNA  
<213> Homo sapien
```

<400> 91
gttcaactcccc cgcctcccca agtcagcagt cctagccccca aaccagccccca gagcagggtc 60

tctctaaagg ggacttgagg gcctgtaaga gcctggtgaa atgggagagt gagaataaaaa 120

tggtctgtga gcagaagctc ctgaagggag agggccccaa gacctcggtgg accagagaac 180

tgaccaacga tggggaaactg atcctgacca tgacggcgga tgacgttgg tgccaccagg 240

tctacgtccg agagtggatg gccacaggtt gaaccggggc cgaagcccac cactggccat 300
gctcaccggcc ctgcttcaact gccccctccg tcccacccccc tccttcttagg atagcgctcc 360

ccttacccca gtcacttctg ggggtcaactg ggatgcctct tgcagggtct tgctttcttt 420
gacctcttct ctcctccctt acaccaacaa agaggaatgg ctgcaagagc ccagatcacc 480

catccgggt tcactccccg cctcccccaag tcagcagtcc tagccccaaa ccagccccaga 540
gcaggggtctc tctaaagggg acttgagggc ctgagcagga aagactggcc ctctagcttc 600

taccctttgt ccctgttagcc tatacagttt agaatattta tttgttaatt ttattaaaat 660
gctttaaaaa aataaaaaaaa aaaaaacaaa aaaaaaaaaag aagagccccgg cgcgcgaaac 720

124

ccgcgtggcc atggcgccggc gaccgcggg gcgcgaaaac agtggctac ctcgcggcct	780
ccccaaattc tccccaccca ccttagcgc agcgaccaac gtgcgcggc cgacagcgaaa	840
gcggccgcga cgagcgccgg acgctacgcg acggacggcg cgggcggca ccacgccacc	900
acgtcacggg cagccgccag cgcacgcccgg gcggcgcgcct gctcacaacc gaggtctgcc	960
tagttgtgc tcccggtgcc gagccaaggc ccgctacgcg cgcacgcgcggc gggctgaggc	1020
agcggcacgc gcgcggcgtg caacgcggc ggcacccggc tggagggggg gaggcacgc	1080
aacacggccg acgcggcgaa gagcgggaac aaacgcacac gacccacacc gcaacggta	1140
gcaacgaccgcg agcggccagc ggacgcgcgc gcgtggcagc aggacgcac gccacgagac	1200
gcgcgagagc gagagaccac tccgaggcgc cggccgggt gtgccaggcc cgacgcgtgg	1260
tggcc	1265

<210> 92
<211> 1406
<212> DNA
<213> Homo sapien

<400> 92	
gattcaagtgc tggcgtttgc gtccgcgttcc ccatccactt actagcgcag gagaaggcta	60
tctcggtccc cagagaagcc tggaccacca cgccggctag atccagagaa cctgacgacc	120
cggcgacggc gacgtctttt ttgactaaaa gacagtgtcc agtgctccag cctaggagtc	180
tacggggacc gcctcccgcg cggccaccat gccaacttc tctggcaact ggaaaatcat	240
ccgatcgaa aacttcgagg aattgctcaa agtgctgggg gtgaatgtga tgctgaggaa	300
gattgtgtgc gtcgcagcgt ccaagccagc agtggagatc aaacaggagg gagacacttt	360
ctacatcaaa acctccacca ccgtgcgcac cacagagatt aacttcaagg ttggggagga	420
gtttgaggag cagactgtgg atggaggccc ctgtaaagcac tgccccctcc gtcccaaaaa	480
ctcccttctag gatagcgctc cccttacccc agtcaatttct ggggtcaact gggatgcctc	540
ttgcagggtc ttgtttctt tgaccttttc tctccccc tacaccaaca aagaggaatg	600
gctgcaagag cccagatcac ccattccggg ttcactcccc gcctcccaa gtcagcgtc	660
ctagccccaa accagccag agcagggtct ctctaaaggg gacttgaggg cctgagcagg	720
aaagactggc cctctagctt ctaccctttg tccctgttagc ctatacagtt tagaatattt	780
atttgttaat ttataaaaa tgctttaaaa aaataaaaaa aaaaaaaca aaaaaaaaaa	840
gaagagcccg gcgcgcgaaa cccgcgtggc catggcgccgg cgacccgcgg ggcgcgaaaa	900
cagtggcgta cctcgccggcc tccccaaattt ctccccaccc acctttagcg cagcgaccaa	960
cgtgcgcgcgc ggcgcagcgaaa ggccggccgcg acgagcgccg gacgctacgc gacggacggc	1020

125

gcgggcccggc accacgccac cacgtcacgg gcagccgcca ggcacgccc	1080
tgctcacaac cgaggctctgc ctagttgctg ctccccgtgc cgagccaagg cccgctacgc	1140
acgcccacgc agggctgagg cagcggcacg cgcgccgt gcaacgccgg cggcacccgg	1200
ctggaggggg ggagggcaccg caacacggcc gacgcggcga agagcggaa caaacgcaca	1260
cgaccacac acgcaacggtg agcaacgacc gagcggccag cggcgaccgc ggcgtggcag	1320
cagggcacga cgccacgaga cgcgagagag cgagagacca ctccgaggcg ccggcccccgg	1380
tgtgccaggc ccgacgcgtg gtggcc	1406

<210> 93
<211> 1441
<212> DNA
<213> Homo sapien

<400> 93	
ccctctctga gtacggagtgttccactgg atccagttca gggttcaatg gagctaggc	60
cagctacggc tcaagatctg gggtccgcct gcgggtgggg tcgcccaggta tccggcacca	120
aggagttgaa tgcaccgagt cagaacctga cgacccggcg acggcgacgt ctctttgac	180
taaaagacag tgtccagtgc tccagcctag gagtctacgg ggaccgcctc ccgcgcgcgc	240
accatgcccacttctctgg caactgaaa atcatccgat cggaaaactt cgaggaattt	300
ctcaaagtgc tgggggtgaa tgtgtatgctg aggaagattt ctgtggctgc agcgtccaag	360
ccagcagtgg agatcaaaca ggagggagac actttctaca tcaaaacctc caccaccgtg	420
cgcaccacag agattaactt caaggttggg gaggagtttggaggcagac tgtggatggg	480
aggccctgtaa cgcactgccc cctccgtccc accccctctc tctaggatag cgctccccctt	540
accccagtca cttctgggg tcactggat gcctcttgc gggctttgtt ttctttgacc	600
tcttcctcc tcccctacac caacaagag gaatggctgc aagagcccgat atcacccatt	660
ccgggttacac tccccgcctc cccaaagtca cagtccttagc cccaaaccag cccagagcag	720
ggtctctcta aaggggactt gagggcctga gcaggaaaga ctggccctct agcttctacc	780
ctttgtccct gttagcctata cagtttagaa tattttttt ttaattttat taaaatgttt	840
aaaaaaaaata aaaaaaaaaa aacaaaaaaaaa aaaaaagaaga gcccggcgcg cgaaacccgc	900
gtggccatgg cgccggcgacc cgccggcgcc gaaaacagtgc gcttacctcg cggccctcccc	960
aaattctccc cacccacctt tagcgcagcg accaacgtgc ggcgcgcga ggggggggg	1020
ccgcgcacgag cgccggacgc tacgcgcacgg acggcgccgg ccggcaccac gccaccacgt	1080
cacgggcagc cgccagcgca cgcccgccgg ggcgcctgctc acaaccgagg tctgccttagt	1140
tgctgctccc ggtgccgagc caaggcccgcc tacgcacgcc cacgcaggc tgaggcagcg	1200

126

gcacgcgcgc	ggcgtgcaac	gccggcgca	cccgctgga	gggggggagg	caccgcaaca	1260
cggccgacgc	ggcgaagagc	gggaacaaac	gcacacgacc	cacaccgcaa	cggtgagcaa	1320
cgaccgagcg	gccagcggcg	accgcggcgt	ggcagcaggc	gacgacgcca	cgagacgcgc	1380
gagagcgaga	gaccactccg	aggcgccggc	ccgggtgtgc	caggcccac	gcgtggtggc	1440
c						1441

<210> 94
<211> 1062
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (19)..(19)
<223> n=a, c, g or t

<220>
<221> misc_feature
<222> (63)..(63)
<223> n=a, c, g or t

<400> 94	gttgaaaaag	gttgggggnc	ccccaaaccc	aagggggtt	aaaggaaaaa	accccccgg	60
	gcnccccggg	gcccggaaaaa	agcccaccac	tggccatgct	caccgcctg	tttcactgcc	120
	ccctccgtcc	cacccctcc	ttcttaggata	gcgcctccct	taccccagtc	acttctgggg	180
	gtcaactggga	tgcctcttgc	agggtcttgc	tttctttgac	ctttctctc	ctccccctaca	240
	ccaaacaaaga	ggaatggctg	caagagccca	gatcacccat	tccgggttca	ctccccgcct	300
	ccccaaagtca	gcagtccctag	ccccaaacca	gcccagagca	gggtctctct	aaaggggact	360
	tgagggcctg	agcaggaaag	actggccctc	tagcttctac	ccttgcctcc	tgtagccatat	420
	acagtttaga	atatttattt	gttaatttta	ttaaatgct	ttaaaaaaat	aaaaaaaaaa	480
	aaacaaaaaa	aaaaaagaag	agccggcgc	gcgaaaacccg	cgtggccatg	gcgcggcgcac	540
	ccgcggggcg	cgaaaaacagt	ggcgtacctc	gcggctccccc	caaattctcc	ccacccacct	600
	ttagcgcagc	gaccaacgtg	cgcgccgcgc	agcgggggchg	gccgcgacga	gcgcggacg	660
	ctacgcgacg	gacggcgcgg	gccggcacca	cgccaccacg	tcacggcag	ccgcccagcgc	720
	acgccccggc	gggcgcctgct	cacaaccgag	gtctgcctag	ttgctgcctcc	cggtggcag	780
	ccaaaggcccc	ctacgcacgc	ccacgcaggg	ctgaggcagc	ggcacgcgcg	cggcgtgcaa	840
	cgccggcggc	acccggctgg	agggggggag	gcaccgcaac	acggccgacg	cggcgaagag	900
	cgggaaacaaa	cgcacacgac	ccacaccgca	acggtgagca	acgaccgagc	ggccagcggc	960

127

gaccgcggcg tggcagcagg cgacgacgcc acgagacgca cgagagcgag agaccactcc 1020
 gagggcgccgg cccgggtgtg ccaggcccga cgcgtggtgg cc 1062

<210> 95
 <211> 937
 <212> DNA
 <213> Homo sapien

<400> 95
 gcggcgccag tgtatggat gcccggcccc gggcaggtcc cagtcacttc tgggggtcac 60
 tggatgcct ctgcagggt ttgcattttt ttgacccctt ctctccccc ctacaccaac 120
 aaagaggaat ggctgcaaga gcccagatca cccattccgg gttcaactccc cgccctcccc 180
 agtcagcagt cctagccca aaccagccca gagcagggtc tctctaaagg ggacttgagg 240
 gcctgagcag gaaagactgg ccctctagct tctaccctt gtccctgttag cctatacagt 300
 ttagaatatt tatttgttaa ttttattaaa atgcttaaaa aaaataaaaa aaaaaaaaaaca 360
 aaaaaaaaaa agaagagccc ggccgcgcaa acccgctgg ccatggcgcc ggcaccccg 420
 gggcgcgaaa acagtggcgt acctcgccgc ctccccaaat tctcccccacc cacctttagc 480
 gcagcgacca acgtgcgcgc cgccgcagg gggcggccgc gacgagccg ggacgctacg 540
 cgacggacgg cgccggccgg caccacgcca ccacgtcactg ggcagccgccc agcgcacgccc 600
 cggcggccgc ctgctcacaa ccgaggtctg cctagttgct gctcccggtg ccgagccaaag 660
 gcccgtacg cacgccccacg caggctgag gcagcggcac gcgcgcggcg tgcaacgcgg 720
 gcccgcacccg gctggaggggg gggaggcacc gcaacacggc cgacgcggcg aagagcggga 780
 acaaacgcac acgacccaca ccgcaacgggt gagcaacgac cgagcggcca gcccgcacccg 840
 cggcgtggca gcaggcgcacg acgccacgag acgcgcgaga gcgagagacc actccgaggc 900
 gccggcccccgt gtgtgccagg cccgacgcgt ggtggcc 937

<210> 96
 <211> 117
 <212> PRT
 <213> Homo sapien

<400> 96

Met	Trp	Thr	Asn	Phe	Gln	Asn	Tyr	Pro	Leu	Cys	Phe	Leu	Gly	Arg	Phe
1				5					10				15		

Arg	Ser	Leu	Thr	Thr	Ala	Phe	Phe	Arg	Asp	Ala	Met	Gly	Phe	Leu	Leu
20					25						30				

Met	Phe	Asp	Leu	Thr	Ser	Gln	Gln	Ser	Phe	Leu	Asn	Val	Arg	Asn	Trp
35					40						45				

Met Ser Gln Leu Gln Ala Asn Ala Tyr Cys Glu Asn Pro Asp Ile Val
50 55 60

Leu Ile Gly Asn Lys Ala Asp Leu Pro Asp Gln Arg Glu Val Asn Glu
65 70 75 80

Arg Gln Ala Arg Glu Leu Ala Asp Lys Tyr Gly Cys Lys Leu Ser Thr
85 90 95

Leu Gly Ile Asn Lys Phe Asp Glu Ala Cys Leu Ser Leu His Gln Trp
100 105 110

Ser Glu Cys Ser Ser
115

<210> 97
<211> 651
<212> PRT
<213> Homo sapien

<400> 97

Met Ala Thr Ala Ser Pro Arg Ser Asp Thr Ser Asn Asn His Ser Gly
1 5 10 15

Arg Leu Gln Leu Gln Val Thr Val Ser Ser Ala Lys Leu Lys Arg Lys
20 25 30

Lys Asn Trp Phe Gly Thr Ala Ile Tyr Thr Glu Val Val Val Asp Gly
35 40 45

Glu Ile Thr Lys Thr Ala Lys Ser Ser Ser Ser Asn Pro Lys Trp
50 55 60

Asp Glu Gln Leu Thr Val Asn Val Thr Pro Gln Thr Thr Leu Glu Phe
65 70 75 80

Gln Val Trp Ser His Arg Thr Leu Lys Ala Asp Ala Leu Leu Gly Lys
85 90 95

Ala Thr Ile Asp Leu Lys Gln Ala Leu Leu Ile His Asn Arg Lys Leu
100 105 110

Glu Arg Val Lys Glu Gln Leu Lys Leu Ser Leu Glu Asn Lys Asn Gly
115 120 125

129

Ile Ala Gln Thr Gly Glu Leu Thr Val Val Leu Asp Gly Leu Val Ile
130 135 140

Glu Gln Glu Asn Ile Thr Asn Cys Ser Ser Ser Pro Thr Ile Glu Ile
145 150 155 160

Gln Glu Asn Gly Asp Ala Leu His Glu Asn Gly Glu Pro Ser Ala Arg
165 170 175

Thr Thr Ala Arg Leu Ala Val Glu Gly Thr Asn Gly Ile Asp Asn His
180 185 190

Val Pro Thr Ser Thr Leu Val Gln Asn Ser Cys Cys Ser Tyr Val Val
195 200 205

Asn Gly Asp Asn Thr Pro Ser Ser Pro Ser Gln Val Ala Ala Arg Pro
210 215 220

Lys Asn Thr Pro Ala Pro Lys Pro Leu Ala Ser Glu Pro Ala Asp Asp
225 230 235 240

Thr Val Asn Gly Glu Ser Ser Ser Phe Ala Pro Thr Asp Asn Ala Ser
245 250 255

Val Thr Gly Thr Pro Val Val Ser Glu Glu Asn Ala Leu Ser Pro Asn
260 265 270

Cys Thr Ser Thr Thr Val Glu Asp Pro Pro Val Gln Glu Ile Leu Thr
275 280 285

Ser Ser Glu Asn Asn Glu Cys Ile Pro Ser Thr Ser Ala Glu Leu Glu
290 295 300

Ser Glu Ala Arg Ser Ile Leu Glu Pro Asp Thr Ser Asn Ser Arg Ser
305 310 315 320

Ser Ser Ala Phe Glu Ala Ala Lys Ser Arg Gln Pro Asp Gly Cys Met
325 330 335

Asp Pro Val Arg Gln Gln Ser Gly Asn Ala Asn Thr Glu Thr Leu Pro
340 345 350

Ser Gly Trp Glu Gln Arg Lys Asp Pro His Gly Arg Thr Tyr Tyr Val
355 360 365

Asp His Asn Thr Arg Thr Thr Trp Glu Arg Pro Gln Pro Leu Pro

130

370 375 380

Pro Gly Trp Glu Arg Arg Val Asp Asp Arg Arg Arg Val Tyr Tyr Val
385 390 395 400

Asp His Asn Thr Arg Thr Thr Trp Gln Arg Pro Thr Met Glu Ser
405 410 415

Val Arg Asn Phe Glu Gln Trp Gln Ser Gln Arg Asn Gln Leu Gln Gly
420 425 430

Ala Met Gln Gln Phe Asn Gln Arg Tyr Leu Tyr Ser Ala Ser Met Leu
435 440 445

Ala Ala Glu Asn Asp Pro Tyr Gly Pro Leu Pro Pro Gly Trp Glu Lys
450 455 460

Arg Val Asp Ser Thr Asp Arg Val Tyr Phe Val Asn His Asn Thr Lys
465 470 475 480

Thr Thr Gln Trp Glu Asp Pro Arg Thr Gln Gly Leu Gln Asn Glu Glu
485 490 495

Thr Leu Gly Arg Arg Leu Arg Gln Phe Arg Ile Phe Ser Val Lys Val
500 505 510

Leu Arg Ser Pro Cys Cys Thr His Ser Thr Gln Gln Pro Thr Pro Phe
515 520 525

Pro Arg Leu Leu Arg Met Arg Lys Pro Thr Asp Thr Ser Asn Gly Gly
530 535 540

Pro Ala Asn Cys Pro Thr Glu Arg Arg Leu Gln Val Lys Pro Ala Lys
545 550 555 560

Tyr Pro Lys Met Gly Pro Ser Leu Met Ala Tyr Pro Arg Thr Gly Thr
565 570 575

Asn Thr Ala Ser Pro Gly Gln Gln Ser Ala Thr Glu Pro Pro Pro Thr
580 585 590

Lys Met Gly Gln Thr Pro Gln Asp Arg Glu Gly Arg His Arg Asn Leu
595 600 605

Thr Ala Glu Pro Ser Thr Asn Gln Gly Thr Arg Lys Glu Pro Pro His
610 615 620

Asn Val Pro Pro Thr Val Gln Thr His Asn Gln Leu Ser Asn Asp Asn
 625 630 635 640

Asn Thr Asn Thr Ile Arg Asn Asn Thr Ser Asn
 645 650

<210> 98
<211> 645
<212> PRT
<213> Homo sapien

<400> 98

Tyr Ile Val Leu Ala Glu Phe Trp Asp Met Ala Thr Ala Ser Pro Arg
 1 5 10 15

Ser Asp Thr Ser Asn Asn His Ser Gly Arg Leu Gln Leu Gln Val Thr
 20 25 30

Val Ser Ser Ala Lys Leu Lys Arg Lys Lys Asn Trp Phe Gly Thr Ala
 35 40 45

Ile Tyr Thr Glu Val Val Val Asp Gly Glu Ile Thr Lys Thr Ala Lys
 50 55 60

Ser Ser Ser Ser Asn Pro Lys Trp Asp Glu Gln Leu Thr Val Asn
 65 70 75 80

Val Thr Pro Gln Thr Thr Leu Glu Phe Gln Val Trp Ser His Arg Thr
 85 90 95

Leu Lys Ala Asp Ala Leu Leu Gly Lys Ala Thr Ile Asp Leu Lys Gln
 100 105 110

Ala Leu Leu Ile His Asn Arg Lys Leu Glu Arg Val Lys Glu Gln Leu
 115 120 125

Lys Leu Ser Leu Glu Asn Lys Asn Gly Ile Ala Gln Thr Gly Glu Leu
 130 135 140

Thr Val Val Leu Asp Gly Leu Val Ile Glu Gln Glu Asn Ile Thr Asn
 145 150 155 160

Cys Ser Ser Ser Pro Thr Ile Glu Ile Gln Glu Asn Gly Asp Ala Leu
 165 170 175

132

His Glu Asn Gly Glu Pro Ser Ala Arg Thr Thr Ala Arg Leu Ala Val
180 185 190

Glu Gly Thr Asn Gly Ile Asp Asn His Val Pro Thr Ser Thr Leu Val
195 200 205

Gln Asn Ser Cys Cys Ser Tyr Val Val Asn Gly Asp Asn Thr Pro Ser
210 215 220

Ser Pro Ser Gln Val Ala Ala Arg Pro Lys Asn Thr Pro Ala Pro Lys
225 230 235 240

Pro Leu Ala Ser Glu Pro Ala Asp Asp Thr Val Asn Gly Glu Ser Ser
245 250 255

Ser Phe Ala Pro Thr Asp Asn Ala Ser Val Thr Gly Thr Pro Val Val
260 265 270

Ser Glu Glu Asn Ala Leu Ser Pro Asn Cys Thr Ser Thr Thr Val Glu
275 280 285

Asp Pro Pro Val Gln Glu Ile Leu Thr Ser Ser Glu Asn Asn Glu Cys
290 295 300

Ile Pro Ser Thr Ser Ala Glu Leu Glu Ser Glu Ala Arg Ser Ile Leu
305 310 315 320

Glu Pro Asp Thr Ser Asn Ser Arg Ser Ser Ala Phe Glu Ala Ala
325 330 335

Lys Ser Arg Gln Pro Asp Gly Cys Met Asp Pro Val Arg Gln Gln Ser
340 345 350

Gly Asn Ala Asn Thr Glu Thr Leu Pro Ser Gly Trp Glu Gln Arg Lys
355 360 365

Asp Pro His Gly Arg Thr Tyr Tyr Val Asp His Asn Thr Arg Thr Thr
370 375 380

Thr Trp Glu Arg Pro Gln Pro Leu Pro Pro Gly Trp Glu Arg Arg Val
385 390 395 400

Asp Asp Arg Arg Arg Val Tyr Tyr Val Asp His Asn Thr Arg Thr Thr
405 410 415

Thr Trp Gln Arg Pro Thr Met Glu Ser Val Arg Asn Phe Glu Gln Trp

133

420	425	430
Gln Ser Gln Arg Asn Gln Leu Gln Gly Ala Met Gln Gln Phe Asn Gln		
435	440	445
Arg Tyr Leu Tyr Ser Ala Ser Met Leu Ala Ala Glu Asn Asp Pro Tyr		
450	455	460
Gly Pro Leu Pro Pro Gly Trp Glu Lys Arg Val Asp Ser Thr Asp Arg		
465	470	475
Val Tyr Phe Val Asn His Asn Thr Lys Thr Thr Gln Trp Glu Asp Pro		
485	490	495
Arg Thr Gln Gly Leu Gln Asn Glu Glu Thr Leu Gly Arg Arg Leu Arg		
500	505	510
Gln Phe Arg Ile Phe Ser Val Lys Val Leu Arg Ser Pro Cys Cys Thr		
515	520	525
His Ser Thr Gln Gln Pro Thr Pro Phe Pro Arg Leu Leu Arg Met Arg		
530	535	540
Lys Pro Thr Asp Thr Ser Asn Gly Gly Pro Ala Asn Cys Pro Thr Glu		
545	550	560
Arg Arg Leu Gln Val Lys Pro Ala Lys Tyr Pro Lys Met Gly Pro Ser		
565	570	575
Leu Met Ala Tyr Pro Arg Thr Gly Thr Asn Thr Ala Ser Pro Gly Gln		
580	585	590
Gln Ser Ala Thr Glu Pro Pro Pro Thr Lys Met Gly Gln Thr Pro Gln		
595	600	605
Asp Arg Glu Gly Arg His Arg Asn Leu Thr Ala Glu Pro Ser Thr Asn		
610	615	620
Gln Gly Thr Arg Lys Glu Pro Thr Pro Gln Arg Thr Thr His Ser Ala		
625	630	635
Asp Ala Gln Pro Thr		
645		

134

<212> PRT

<213> Homo sapien

<400> 99

Met Gly Pro Gly Gly Pro Leu Leu Ser Pro Ser Arg Gly Phe Leu Leu
1 5 10 15

Cys Lys Thr Gly Trp His Ser Asn Arg Leu Leu Gly Asp Cys Gly Pro
20 25 30

His Thr Pro Val Ser Thr Ala Leu Ser Phe Ile Ala Val Gly Met Ala
35 40 45

Ala Pro Ser Met Lys Glu Arg Gln Val Cys Trp Gly Ala Arg Asp Glu
50 55 60

Tyr Trp Lys Cys Leu Asp Glu Asn Leu Glu Asp Ala Ser Gln Cys Lys
65 70 75 80

Lys Leu Arg Ser Ser Phe Glu Ser Ser Cys Pro Gln Gln Trp Ile Lys
85 90 95

Tyr Phe Asp Lys Arg Arg Asp Tyr Leu Lys Phe Lys Glu Lys Phe Glu
100 105 110

Ala Gly Gln Phe Glu Pro Ser Glu Thr Thr Ala Lys Ser
115 120 125

<210> 100

<211> 164

<212> PRT

<213> Homo sapien

<400> 100

Phe Phe Leu Glu Pro Cys Ala Pro Leu Leu Ala Glu Pro Leu Leu Glu
1 5 10 15

Arg Asp Glu Ala Glu Gly Val Gly Gly Ala Asp Ala Gly Pro Ala Leu
20 25 30

Leu Tyr Gly Leu Val Gly Asp Gly Glu Leu Ala Gln Val Val Ala Asn
35 40 45

His Leu Gly Leu Asp Leu His Leu Val Glu Gly Leu Ala Val Val Asp
50 55 60

Ala His His Ala Ala His His Leu Gly Gln Asp Asp His Val Pro Gln

135

65	70	75	80
----	----	----	----

Val Arg Leu His His Phe Arg Leu Leu His Gly Arg Arg Leu Leu Leu
85 90 95

Gly Leu Ala Gln Ala Leu Gln Gln Gly Val Leu Leu Pro Pro Gln Ala
100 105 110

Pro Val Gln Pro Pro Pro Leu Ala Arg Thr Val Gln Leu His Gln Leu
115 120 125

Leu Val Gly His Val Gln Gln Leu Val Glu Val His Ala Ala Ala Leu His
130 135 140

Arg Ser Arg Asn Gly Ser Pro Ile Tyr Glu Gly Lys Thr Gly Leu Leu
 145 150 155 160

Gly Gly Pro Gly

<210> 101
<211> 129
<212> PRT
<213> Homo sapien

<400> 101

Arg Asp Glu Ala Glu Gly Val Gly Gly Ala Asp Ala Gly Pro Ala Leu
20 25 30

Leu Tyr Gly Leu Val Gly Asp Gly Glu Leu Ala Gln Val Val Ala Asn
35 40 45

His Leu Gly Leu Asp Leu His Leu Val Glu Gly Leu Ala Val Val Asp
50 55 60

Ala His His Ala Ala His His Leu Gly Gln Asp Asp His Val Pro Gln
65 70 75 80

Val Arg Leu His His Phe Arg Leu Leu His Gly Arg Arg Leu Leu Leu
85 90 95

Gly Leu Ala Gln Ala Leu Gln Gln Gly Val Leu Leu Pro Pro Gln Ala
100 105 110

Pro Val Gln Pro Pro Arg Trp Arg Ala Leu Tyr Ser Cys Ile Ser Cys
 115 120 125

Ser

<210> 102
 <211> 139
 <212> PRT
 <213> Homo sapien

<400> 102

Asp Pro Arg Trp Ala Leu Tyr Ser Leu Tyr Val Tyr Lys Phe Leu His
 1 5 10 15

Phe Ser Tyr Ser Ser Ala Lys Asn Pro Asp Gly Cys Phe Phe Gln Lys
 20 25 30

Val Leu Asn Gly Phe Thr Lys Phe Phe Cys Lys Glu Gln Tyr Cys Lys
 35 40 45

Leu Leu Lys Leu Tyr Phe Tyr Arg Leu Phe Ala Leu Leu Trp Ile Leu
 50 55 60

Cys Leu Ser Gly Phe Leu Lys Phe Phe Phe Tyr Ser Glu Ile Met Glu
 65 70 75 80

Leu Val Leu Ala Ala Gly Ala Leu Leu Phe Cys Gly Phe Ile Ile
 85 90 95

Tyr Asp Thr His Ser Leu Met His Lys Leu Ser Pro Glu Glu Tyr Val
 100 105 110

Leu Ala Ala Ile Ser Leu Tyr Leu Asp Ile Ile Asn Leu Phe Leu His
 115 120 125

Leu Leu Arg Phe Leu Glu Ala Val Asn Lys Lys
 130 135

<210> 103
 <211> 525
 <212> PRT
 <213> Homo sapien

<400> 103

Met Gly Asp Leu Glu Leu Leu Leu Pro Gly Glu Ala Glu Val Leu Val
 1 5 10 15

Arg Gly Leu Arg Ser Phe Pro Leu Arg Glu Met Gly Ser Glu Gly Trp
20 25 30

Asn Gln Gln His Glu Asn Leu Glu Lys Leu Asn Met Gln Ala Ile Leu
35 40 45

Asp Ala Thr Val Ser Gln Gly Glu Pro Ile Gln Glu Leu Leu Val Thr
50 55 60

His Gly Lys Val Pro Thr Leu Val Glu Glu Leu Ile Ala Val Glu Met
65 70 75 80

Trp Lys Gln Lys Val Phe Pro Val Phe Cys Arg Val Glu Asp Phe Lys
85 90 95

Pro Gln Asn Thr Phe Pro Ile Tyr Met Val Val His His Glu Ala Ser
100 105 110

Ile Ile Asn Leu Leu Glu Thr Val Phe Phe His Lys Glu Val Cys Glu
115 120 125

Ser Ala Glu Asp Thr Val Leu Asp Leu Val Asp Tyr Cys His Arg Lys
130 135 140

Leu Thr Leu Leu Val Ala Gln Ser Gly Cys Gly Gly Pro Pro Glu Gly
145 150 155 160

Glu Gly Ser Gln Asp Ser Asn Pro Met Gln Glu Leu Gln Lys Gln Ala
165 170 175

Glu Leu Met Glu Phe Glu Ile Ala Leu Lys Ala Leu Ser Val Leu Arg
180 185 190

Tyr Ile Thr Asp Cys Val Asp Ser Leu Ser Leu Ser Thr Leu Ser Arg
195 200 205

Met Leu Ser Thr His Asn Leu Pro Cys Leu Leu Val Glu Leu Leu Glu
210 215 220

His Ser Pro Trp Ser Arg Arg Glu Gly Gly Lys Leu Gln Gln Phe Glu
225 230 235 240

Gly Ser Arg Trp His Thr Val Ala Pro Ser Glu Gln Gln Lys Leu Ser
245 250 255

Lys Leu Asp Gly Gln Val Trp Ile Ala Leu Tyr Asn Leu Leu Ser
260 265 270

Pro Glu Ala Gln Ala Arg Tyr Cys Leu Thr Ser Phe Ala Lys Gly Arg
275 280 285

Leu Leu Lys Val Arg Leu Pro Pro His Gln Pro Pro Gln Pro Gln Tyr
290 295 300

Arg Pro Pro His Pro Thr Pro Thr Ala Ser Leu Leu Phe Ile Phe Ala
305 310 315 320

His Pro Pro Gln Pro Gln Cys Ser Phe Gln Ser Leu Gly Leu Ser Asp
325 330 335

Thr Pro Ala Ser Gly Thr Trp Ala Pro Thr Gly Ile Leu Ser Pro Thr
340 345 350

Gln Pro Leu Pro Phe Pro Trp Pro Pro Gly Gln His Leu His His Thr
355 360 365

Gly Leu His Trp Thr Pro Leu Gln Leu Arg Ala Phe Leu Thr Asp Thr
370 375 380

Leu Leu Asp Gln Leu Pro Asn Leu Ala His Leu Gln Ser Phe Leu Ala
385 390 395 400

His Leu Thr Leu Thr Glu Thr Gln Pro Pro Lys Lys Asp Leu Val Leu
405 410 415

Glu Gln Ile Pro Glu Ile Trp Glu Arg Leu Glu Arg Glu Asn Arg Gly
420 425 430

Lys Trp Gln Ala Ile Ala Lys His Gln Leu Gln His Val Phe Ser Pro
435 440 445

Ser Glu Gln Asp Leu Arg Leu Gln Ala Arg Arg Trp Ala Glu Thr Tyr
450 455 460

Arg Leu Asp Val Leu Glu Ala Val Ala Pro Glu Arg Pro Arg Cys Ala
465 470 475 480

Tyr Cys Ser Ala Glu Ala Ser Lys Arg Cys Ser Arg Cys Gln Asn Glu
485 490 495

139

Trp Tyr Cys Cys Arg Glu Cys Gln Val Lys His Trp Glu Lys His Gly
500 505 510

Lys Thr Cys Val Leu Ala Ala Gln Gly Asp Arg Ala Lys
515 520 525

<210> 104
<211> 385
<212> PRT
<213> Homo sapien

<400> 104

Pro Phe Pro Trp Leu Arg Glu Leu Thr Leu Pro Asn Arg Pro Ala Thr
1 5 10 15

Val Leu Ser Gln Thr Leu Ala Pro Ser Gly Ser Val Val Pro Glu Cys
20 25 30

Asp Ser Ile Pro Thr Pro Ala Ala Gln Asp Pro Pro Asp Pro Gly
35 40 45

Leu Asp Met Gly Asp Leu Glu Leu Leu Leu Pro Gly Glu Ala Glu Val
50 55 60

Leu Val Arg Gly Leu Arg Ser Phe Pro Leu Arg Glu Met Gly Ser Glu
65 70 75 80

Gly Trp Asn Gln Gln His Glu Asn Leu Glu Lys Leu Asn Met Gln Ala
85 90 95

Ile Leu Asp Ala Thr Val Ser Gln Gly Glu Pro Ile Gln Glu Leu Leu
100 105 110

Val Thr His Gly Lys Val Pro Thr Leu Val Glu Glu Leu Ile Ala Val
115 120 125

Glu Met Trp Lys Gln Lys Val Phe Pro Val Phe Cys Arg Val Glu Asp
130 135 140

Phe Lys Pro Gln Asn Thr Phe Pro Ile Tyr Met Val Val His His Glu
145 150 155 160

Ala Ser Ile Ile Asn Leu Leu Glu Thr Val Phe Phe His Lys Glu Val
165 170 175

Cys Glu Ser Ala Glu Asp Thr Val Leu Asp Leu Val Asp Tyr Cys His
180 185 190

140

Arg Lys Leu Thr Leu Leu Val Ala Gln Ser Gly Cys Gly Gly Pro Pro
195 200 205

Glu Gly Glu Gly Ser Gln Asp Ser Asn Pro Met Gln Glu Leu Gln Lys
210 215 220

Gln Ala Glu Leu Met Glu Phe Glu Ile Ala Leu Lys Ala Leu Ser Val
225 230 235 240

Leu Arg Tyr Ile Thr Asp Cys Val Asp Ser Leu Ser Leu Ser Thr Leu
245 250 255

Ser Arg Met Leu Ser Thr His Asn Leu Pro Cys Leu Leu Val Glu Leu
260 265 270

Leu Glu His Ser Pro Trp Ser Arg Arg Glu Gly Gly Lys Leu Gln Gln
275 280 285

Phe Glu Gly Ser Arg Trp His Thr Val Ala Pro Ser Glu Gln Gln Lys
290 295 300

Leu Ser Lys Leu Asp Gly Gln Val Trp Ile Ala Leu Tyr Asn Leu Leu
305 310 315 320

Leu Ser Pro Glu Ala Gln Ala Arg Tyr Cys Leu Thr Ser Phe Ala Lys
325 330 335

Gly Arg Leu Leu Lys Val Arg Leu Pro Pro His Gln Pro Pro Gln Pro
340 345 350

Gln Tyr Arg Pro Pro His Pro Thr Pro Thr Ala Ser Leu Leu Phe Ile
355 360 365

Phe Ala His Pro Pro Gln Pro Gln Cys Ser Phe Gln Ser Leu Gly Leu
370 375 380

Arg
385

<210> 105
<211> 438
<212> PRT
<213> Homo sapien

<400> 105

141

Met Asp Glu Ile Glu Lys Tyr Gln Glu Val Glu Asp Gln Asp Pro
1 5 10 15

Ser Cys Pro Arg Leu Ser Arg Glu Leu Leu Asp Glu Lys Glu Pro Glu
20 25 30

Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser Gly Tyr
35 40 45

Leu Glu Leu Pro Asp Leu Gly Gln Pro Tyr Ser Ser Ala Val Tyr Ser
50 55 60

Leu Glu Glu Gln Tyr Leu Gly Leu Ala Leu Asp Val Asp Arg Ile Lys
65 70 75 80

Lys Asp Gln Glu Glu Glu Asp Gln Gly Pro Pro Cys Pro Arg Leu
85 90 95

Ser Arg Glu Leu Leu Glu Val Val Glu Pro Glu Val Leu Gln Asp Ser
100 105 110

Leu Asp Arg Cys Tyr Ser Thr Pro Ser Ser Cys Leu Glu Gln Pro Asp
115 120 125

Ser Cys Gln Pro Tyr Gly Ser Ser Phe Tyr Ala Leu Glu Glu Lys His
130 135 140

Val Gly Phe Ser Leu Asp Val Gly Glu Ile Glu Lys Lys Gly Lys Gly
145 150 155 160

Lys Lys Arg Arg Gly Arg Arg Ser Lys Lys Glu Arg Arg Arg Gly Arg
165 170 175

Lys Glu Gly Glu Glu Asp Gln Asn Pro Pro Cys Pro Arg Leu Ser Arg
180 185 190

Glu Leu Leu Asp Glu Lys Gly Pro Glu Val Leu Gln Asp Ser Leu Asp
195 200 205

Arg Cys Tyr Ser Thr Pro Ser Gly Cys Leu Glu Leu Thr Asp Ser Cys
210 215 220

Gln Pro Tyr Arg Ser Ala Phe Tyr Val Leu Glu Gln Gln Arg Val Gly
225 230 235 240

Leu Ala Val Asp Met Asp Glu Ile Glu Lys Tyr Gln Glu Val Glu Glu

142
245 250 255

Asp Gln Asp Pro Ser Cys Pro Arg Leu Ser Arg Glu Leu Leu Asp Glu
260 265 270

Lys Glu Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr
275 280 285

Pro Ser Gly Tyr Leu Glu Leu Pro Asp Leu Gly Gln Pro Tyr Ser Ser
290 295 300

Ala Val Tyr Ser Leu Glu Glu Gln Tyr Leu Gly Leu Ala Leu Asp Val
305 310 315 320

Asp Lys Ile Glu Lys Lys Gly Lys Lys Arg Arg Gly Arg Arg
325 330 335

Ser Lys Lys Glu Arg Arg Gly Ser Lys Glu Gly Glu Glu Asp Gln
340 345 350

Asn Pro Pro Cys Pro Arg Leu Ser Gly Val Leu Met Glu Val Glu Glu
355 360 365

Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser
370 375 380

Met Tyr Phe Glu Leu Pro Asp Ser Phe Gln His Tyr Arg Ser Val Phe
385 390 395 400

Tyr Ser Phe Glu Glu Gln His Ile Ser Phe Ala Leu Asp Val Asp Asn
405 410 415

Arg Phe Leu Thr Leu Met Gly Thr Ser Leu His Leu Val Phe Gln Met
420 425 430

Gly Val Ile Phe Pro Gln
435

<210> 106
<211> 334
<212> PRT
<213> Homo sapien

<400> 106

Ser Leu Lys Ser Cys Arg Thr His Trp Ile Asp Val Ile Gln Leu Leu
1 5 10 15

Pro Val Val Leu Asn Ser Leu Thr Pro Ala Ser Pro Met Glu Val Pro
20 25 30

Phe Met His Trp Arg Lys Asn Met Leu Ala Phe Leu Leu Thr Trp Glu
35 40 45

Lys Leu Lys Arg Arg Gly Arg Gly Arg Lys Glu Gly Glu Glu Asp Gln
50 55 60

Arg Arg Lys Glu Arg Arg Gly Arg Lys Glu Gly Glu Glu Asp Gln Asn
65 70 75 80

Pro Pro Cys Pro Arg Leu Ser Arg Glu Leu Leu Asp Glu Lys Gly Pro
85 90 95

Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser Gly
100 105 110

Cys Leu Glu Leu Thr Asp Ser Cys Gln Pro Tyr Arg Ser Ala Phe Tyr
115 120 125

Val Leu Glu Gln Gln Arg Val Gly Leu Ala Val Asp Met Asp Glu Ile
130 135 140

Glu Lys Tyr Gln Glu Val Glu Glu Asp Gln Asp Pro Ser Cys Pro Arg
145 150 155 160

Leu Ser Arg Glu Leu Leu Asp Glu Lys Glu Pro Glu Val Leu Gln Asp
165 170 175

Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser Gly Tyr Leu Glu Leu Pro
180 185 190

Asp Leu Gly Gln Pro Tyr Ser Ser Ala Val Tyr Ser Leu Glu Glu Gln
195 200 205

Tyr Leu Gly Leu Ala Leu Asp Val Asp Lys Ile Glu Lys Lys Gly Lys
210 215 220

Gly Lys Lys Arg Arg Gly Arg Arg Ser Lys Lys Glu Arg Arg Arg Gly
225 230 235 240

Ser Lys Glu Gly Glu Glu Asp Gln Asn Pro Pro Cys Pro Arg Leu Ser
245 250 255

144

Gly Val Leu Met Glu Val Glu Glu Pro Glu Val Leu Gln Asp Ser Leu
 260 265 270

Asp Arg Cys Tyr Ser Thr Pro Ser Met Tyr Phe Glu Leu Pro Asp Ser
 275 280 285

Phe Gln His Tyr Arg Ser Val Phe Tyr Ser Phe Glu Glu Gln His Ile
 290 295 300

Ser Phe Ala Leu Asp Val Asp Asn Arg Phe Leu Thr Leu Met Gly Thr
 305 310 315 320

Ser Leu His Leu Val Phe Gln Met Gly Val Ile Phe Pro Gln
 325 330

<210> 107
<211> 140
<212> PRT
<213> Homo sapien

<400> 107

Met Arg Arg Arg Ser His Ser Thr Arg Leu Ser Ala Gly Gly Ser Trp
 1 5 10 15

Ser Pro His His Leu Leu Ser Pro Ser Tyr Ser Val Lys Ser Arg Asp
 20 25 30

Arg Lys Met Val Gly Asp Val Thr Gly Ala Gln Ala Tyr Ala Ser Thr
 35 40 45

Ala Lys Cys Leu Asn Ile Trp Ala Leu Ile Leu Gly Ile Leu Met Thr
 50 55 60

Ile Gly Phe Ile Leu Leu Val Phe Gly Ser Val Thr Val Ser His
 65 70 75 80

Ile Met Phe Gln Asn Asn Thr Gly Lys Thr Gly Leu Leu Val Ala Ala
 85 90 95

His Ser Leu Gln Pro Leu His Ser Thr Val Gln Cys Trp Pro Cys Asn
 100 105 110

Ala Val Ala Val Ala Pro Ala Pro Leu Val Leu Pro Leu Asn Thr Ala
 115 120 125

Val Tyr Thr His Thr Pro Val Tyr Ser Val Ile Gln
 130 135 140

<210> 108
<211> 114
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (53)..(53)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (82)..(82)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (94)..(94)
<223> X=any amino acid

<400> 108

Gly Gln Glu Asp Gly Trp Arg Arg Asp Arg Gly Pro Gly Leu Cys Leu
1 5 10 15

His Arg Gln Val Pro Glu His Leu Gly Pro Asp Ser Gly His Pro His
20 25 30

Asp His Trp Ile His Pro Val Thr Gly Ile Arg Leu Cys Asp Ser Leu
35 40 45

Pro Tyr Tyr Val Xaa Asp Asn Thr Gly Lys Thr Gly Leu Leu Val Ala
50 55 60

Ala His Ser Leu Gln Pro Leu His Ser Thr Val Gln Cys Trp Pro Cys
65 70 75 80

Thr Xaa Gly Cys Cys Pro Cys Pro Leu Gly Pro Ala Pro Xaa Tyr Ser
85 90 95

Ser Leu Tyr Pro His Thr Cys Leu Gln Cys His Ser Ile Lys Arg Thr
100 105 110

Cys Leu

<210> 109
<211> 182

146

<212> PRT

<213> Homo sapien

<400> 109

Met Glu Glu Met Lys Asn Glu Ala Glu Thr Thr Ser Met Val Ser Met
1 5 10 15

Pro Leu Tyr Ala Val Met Tyr Pro Val Phe Asn Glu Leu Glu Arg Val
20 25 30

Asn Leu Ser Ala Ala Gln Thr Leu Arg Ala Ala Phe Ile Lys Ala Glu
35 40 45

Lys Glu Asn Pro Gly Leu Thr Gln Asp Ile Ile Met Lys Ile Leu Glu
50 55 60

Lys Lys Ser Val Glu Val Asn Phe Thr Glu Ser Leu Leu Arg Met Ala
65 70 75 80

Ala Asp Asp Val Glu Glu Tyr Met Ile Glu Arg Pro Glu Pro Glu Phe
85 90 95

Gln Asp Leu Asn Glu Lys Ala Arg Ala Leu Lys Gln Ile Leu Ser Lys
100 105 110

Ile Pro Asp Glu Ile Asn Asp Arg Val Arg Phe Leu Gln Thr Ile Lys
115 120 125

Ala Leu Glu His Gln Lys Lys Glu Phe Val Lys Tyr Ser Lys Ser Phe
130 135 140

Ser Asp Thr Leu Lys Thr Tyr Phe Lys Asp Gly Lys Ala Ile Asn Val
145 150 155 160

Phe Val Ser Ala Asn Arg Leu Ile His Gln Thr Asn Leu Ile Leu Gln
165 170 175

Thr Phe Lys Thr Val Ala
180

<210> 110

<211> 141

<212> PRT

<213> Homo sapien

<400> 110

Met Arg Met Thr Met Glu Glu Met Lys Asn Glu Ala Glu Thr Thr Ser

1	5	10	15	147
				Met Val Ser Met Pro Leu Tyr Ala Val Met Tyr Pro Val Phe Asn Glu
	20	25		30
				Leu Glu Arg Val Asn Leu Ser Ala Ala Gln Thr Leu Arg Ala Ala Phe
	35	40		45
				Ile Lys Ala Glu Lys Glu Asn Pro Gly Leu Thr Gln Asp Ile Ile Met
	50	55		60
				Lys Ile Leu Glu Lys Lys Ser Val Glu Val Asn Phe Thr Glu Ser Leu
	65	70	75	80
				Leu Arg Met Ala Ala Asp Asp Val Glu Glu Tyr Met Ile Glu Arg Pro
	85		90	95
				Glu Pro Glu Phe Gln Asp Leu Asn Glu Lys Ala Arg Ala Leu Lys Gln
	100		105	110
				Ile Leu Ser Lys Ile Pro Asp Glu Ile Asn Asp Arg Val Arg Phe Leu
	115		120	125
				Gln Thr Ile Lys His Leu Asn Thr Lys Arg Lys Asn Leu
	130		135	140
<210>	111			
<211>	132			
<212>	PRT			
<213>	Homo sapien			
<400>	111			
				Gly Arg Val Pro Leu Ala Leu Gly Val Gln Thr Leu Pro Gln Thr Cys
1	5		10	15
				Asp Glu Pro Lys Ala His Thr Ser Phe Gln Ile Ser Leu Ser Val Ser
	20		25	30
				Tyr Thr Gly Ser Ser Gly Arg Pro Gly Arg Tyr Glu Leu Phe Lys Ser
	35		40	45
				Ser Pro His Ser Leu Phe Pro Glu Lys Met Val Ser Ser Cys Leu Asp
	50		55	60
				Ala His Thr Gly Ile Ser His Glu Asp Leu Ile Gln Val Gly Gly Pro
	65		70	75
				80

148

Pro Ile Ser Leu Gln Ile His Asp Ser Pro Ala Leu Ala Ser Ala Ser
85 90 95

Pro Pro Leu Ser Pro Val Pro Pro Leu Tyr Val Val Glu Arg Ala Lys
100 105 110

Ser Gln Ser Cys Val Thr Gly Asp Ser His Phe Pro Cys Leu Ser Ile
115 120 125

Ser Phe Phe Tyr
130

<210> 112
<211> 277
<212> PRT
<213> Homo sapien

<400> 112

Met Glu Leu Asp Leu Ser Pro Pro His Leu Ser Ser Ser Pro Glu Asp
1 5 10 15

Leu Cys Pro Ala Pro Gly Thr Pro Pro Gly Thr Pro Arg Pro Pro Asp
20 25 30

Thr Pro Leu Pro Glu Glu Val Lys Arg Ser Gln Pro Leu Leu Ile Pro
35 40 45

Thr Thr Gly Arg Lys Leu Arg Glu Glu Glu Arg Arg Ala Thr Ser Leu
50 55 60

Pro Ser Ile Pro Asn Pro Phe Pro Glu Leu Cys Ser Pro Pro Ser Gln
65 70 75 80

Ser Pro Ile Leu Gly Gly Pro Ser Ser Ala Arg Gly Leu Leu Pro Arg
85 90 95

Asp Ala Ser Arg Pro His Val Val Lys Val Tyr Ser Glu Asp Gly Ala
100 105 110

Cys Arg Ser Val Glu Val Ala Ala Gly Ala Thr Ala Arg His Val Cys
115 120 125

Glu Met Leu Val Gln Arg Ala His Ala Leu Ser Asp Glu Thr Trp Gly
130 135 140

Leu Val Glu Cys His Pro His Leu Ala Leu Glu Arg Gly Leu Glu Asp

149

145

150

155

160

His Glu Ser Val Val Glu Val Gln Ala Ala Trp Pro Val Gly Gly Asp
165 170 175

Ser Arg Phe Val Phe Arg Lys Asn Phe Ala Lys Tyr Glu Leu Phe Lys
180 185 190

Ser Ser Pro His Ser Leu Phe Pro Glu Lys Met Val Ser Ser Cys Leu
195 200 205

Asp Ala His Thr Gly Ile Ser His Glu Asp Leu Ile Gln Val Gly Gly
210 215 220

Pro Pro Ile Ser Leu Gln Ile His Asp Ser Pro Ala Leu Ala Ser Ala
225 230 235 240

Ser Pro Pro Leu Ser Pro Val Pro Pro Leu Tyr Val Val Glu Arg Ala
245 250 255

Lys Ser Gln Ser Cys Val Thr Gly Asp Ser His Phe Pro Cys Leu Ser
260 265 270

Ile Ser Phe Phe Tyr
275

<210> 113
<211> 155
<212> PRT
<213> Homo sapien

<400> 113

Lys Glu Ser Leu Phe Ala Tyr Leu Gly Leu Ser Pro Pro Ser Tyr Thr
 20 25 30

Phe Pro Ala Pro Ala Ala Val Ile Pro Thr Glu Ala Ala Ile Tyr Gln
35 40 45

Pro Ser Val Ile Leu Asn Pro Arg Ala Leu Gln Pro Ser Thr Ala Tyr
50 55 60

Tyr Pro Ala Gly Thr Gln Leu Phe Met Asn Tyr Thr Ala Tyr Tyr Tyr Pro
65 70 75 80

150

Ser Pro Pro Gly Ser Pro Asn Ser Leu Gly Tyr Phe Pro Thr Ala Ala
85 90 95

Asn Leu Ser Gly Val Pro Pro Gln Pro Gly Thr Val Val Arg Met Gln
100 105 110

Gly Leu Ala Tyr Asn Thr Gly Val Lys Glu Ile Leu Asn Phe Phe Gln
115 120 125

Gly Tyr Gln Tyr Ala Thr Glu Asp Gly Leu Ile His Thr Asn Asp Gln
130 135 140

Ala Arg Thr Leu Pro Lys Glu Trp Val Cys Ile
145 150 155

<210> 114

<211> 103

<212> PRT

<213> Homo sapien

<400> 114

Met Val Lys Leu Asn Ser Asn Pro Ser Glu Lys Gly Thr Lys Pro Pro
1 5 10 15

Ser Val Glu Asp Gly Phe Gln Thr Val Pro Leu Ile Thr Pro Leu Glu
20 25 30

Val Asn His Leu Gln Leu Pro Ala Pro Glu Lys Val Ile Val Lys Thr
35 40 45

Arg Thr Glu Tyr Gln Pro Glu Gln Lys Asn Lys Gly Lys Phe Arg Val
50 55 60

Pro Lys Ile Ala Glu Phe Thr Val Thr Ile Leu Val Ser Leu Ala Leu
65 70 75 80

Ala Phe Leu Ala Cys Ile Val Phe Leu Val Val Tyr Lys Ala Phe Thr
85 90 95

Tyr Leu Lys Glu Leu Asn Ser
100

<210> 115

<211> 117

<212> PRT

<213> Homo sapien

151

```
<220>
<221> MISC_FEATURE
<222> (114)..(114)
<223> X=any amino acid
```

<400> 115

```

Pro Pro Thr Ser Ala Ala Gln Ser Gly Lys Lys Gly Val Arg Met Val
1           5                   10                  15

```

Lys Leu Asn Ser Asn Pro Ser Glu Lys Gly Thr Lys Pro Pro Ser Val
20 25 30

Glu Asp Gly Phe Gln Thr Val Pro Leu Ile Thr Pro Leu Glu Val Asn
 35 40 45

His Leu Gln Leu Pro Ala Pro Glu Lys Val Ile Val Lys Thr Arg Thr
50 55 60

Glu Tyr Gln Pro Glu Gln Lys Asn Lys Gly Lys Phe Arg Val Pro Lys
 65 70 75 80

Ile Ala Glu Phe Thr Val Thr Ile Leu Val Ser Leu Ala Leu Ala Phe
85 90 95

Leu Ala Cys Ile Val Phe Leu Val Val Tyr Lys Ala Phe Thr Tyr Leu
100 105 110

Lys Xaa Leu Asn Ser
115

<210> 116
<211> 454
<212> PRT
<213> Homo sapien

<400> 116

Met Pro Glu Phe Leu Glu Asp Pro Ser Val Leu Thr Lys Asp Lys Leu
1 5 10 15

Lys Ser Glu Leu Val Ala Asn Asn Val Thr Leu Pro Ala Gly Glu Gln
20 25 30

Arg Lys Asp Val Tyr Val Gln Leu Tyr Leu Gln His Leu Thr Ala Arg
35 40 45

Asn Arg Pro Pro Leu Pro Ala Gly Thr Asn Ser Lys Gly Pro Pro Pro Asp
50 55 60

Ala Ala Ala Ala Gly Arg Ser Arg Ala Ala Val Gly Arg Lys Ala Thr
85 90 95

Lys Lys Thr Asp Lys Pro Arg Gln Glu Asp Lys Asp Asp Asp Leu Asp Val
 100 105 110

Thr Glu Leu Thr Asn Glu Asp Leu Leu Asp Gln Leu Val Lys Tyr Gly
115 120 125

Val Asn Pro Gly Pro Ile Val Gly Thr Thr Arg Lys Leu Tyr Glu Lys
 130 135 140

Lys Leu Leu Lys Leu Arg Glu Gln Gly Thr Glu Ser Arg Ser Ser Thr
145 150 155 160

Pro Leu Pro Thr Ile Ser Ser Ser Ala Glu Asn Thr Arg Gln Asn Gly
165 170 175

Ser Asn Asp Ser Asp Arg Tyr Ser Asp Asn Glu Glu Asp Ser Lys Ile
180 185 190

Glu Leu Lys Leu Glu Lys Arg Glu Pro Leu Lys Gly Arg Ala Lys Thr
195 200 205

Pro Val Thr Leu Lys Gln Arg Arg Val Glu His Asn Gln Ser Tyr Ser
210 215 220

Gln Ala Gly Ile Thr Glu Thr Glu Trp Thr Ser Gly Ser Ser Lys Gly
225 230 235 240

Gly Pro Leu Gln Ala Leu Thr Arg Glu Ser Thr Arg Gly Ser Arg Arg
245 250 255

Thr Pro Arg Lys Arg Val Glu Thr Ser Glu His Phe Arg Ile Asp Gly
260 265 270

Pro Val Ile Ser Glu Ser Thr Pro Ile Ala Glu Thr Ile Met Ala Ser
275 280 285

Ser Asn Glu Ser Leu Val Val Asn Arg Val Thr Gly Asn Phe Lys His
280 285 290

153

Ala Ser Pro Ile Leu Pro Ile Thr Glu Phe Ser Asp Ile Pro Arg Arg
305 310 315 320

Ala Pro Lys Lys Pro Leu Thr Arg Ala Glu Val Gly Glu Lys Thr Glu
325 330 335

Glu Arg Arg Val Glu Arg Asp Ile Leu Lys Glu Met Phe Pro Tyr Glu
340 345 350

Ala Ser Thr Pro Thr Gly Ile Ser Ala Ser Cys Arg Arg Pro Ile Lys
355 360 365

Gly Ala Ala Gly Arg Pro Leu Glu Leu Ser Asp Phe Arg Met Glu Glu
370 375 380

Ser Phe Ser Ser Lys Tyr Val Pro Lys Tyr Val Pro Leu Ala Asp Val
385 390 395 400

Lys Ser Glu Lys Thr Lys Lys Gly Arg Ser Ile Pro Val Trp Ile Lys
405 410 415

Ile Leu Leu Phe Val Val Val Ala Val Phe Leu Phe Leu Val Tyr Gln
420 425 430

Ala Met Glu Thr Asn Gln Val Asn Pro Phe Ser Asn Phe Leu His Val
435 440 445

Asp Pro Arg Lys Ser Asn
450

<210> 117
<211> 380
<212> PRT
<213> Homo sapien

<400> 117

Met Glu Leu Gly Arg Pro Leu Leu Glu Val Leu Ala Ser Ala Leu Ser
1 5 10 15

Pro Ala Ser Pro Pro Leu Leu Pro Pro Asp Tyr Ile Leu Cys Val Val
20 25 30

Ser Leu Leu Gln Met Lys Asp Leu Gly Ala Glu His Leu Ala Gly His
35 40 45

Glu Gly Val Gln Leu Leu Gly Leu Leu Asn Val Tyr Leu Glu Gln Glu

50	55	154	60
Glu Arg Phe Gln Pro Arg Glu Lys Gly Leu Ser Leu Ile Glu Ala Thr			
65	70	75	80
Pro Glu Asn Asp Asn Thr Leu Cys Pro Gly Leu Arg Asn Ala Lys Val			
85	90	95	
Glu Asp Leu Arg Ser Leu Ala Asn Phe Phe Gly Ser Cys Thr Glu Thr			
100	105	110	
Phe Val Leu Ala Val Asn Ile Leu Asp Arg Phe Leu Ala Leu Met Lys			
115	120	125	
Val Lys Pro Lys His Leu Ser Cys Ile Gly Val Cys Ser Phe Leu Leu			
130	135	140	
Ala Ala Arg Ile Val Glu Glu Asp Cys Asn Ile Pro Ser Thr His Asp			
145	150	155	160
Val Ile Arg Ile Ser Gln Cys Lys Cys Thr Ala Ser Asp Ile Lys Arg			
165	170	175	
Met Glu Lys Ile Ile Ser Glu Lys Leu His Tyr Glu Leu Glu Ala Thr			
180	185	190	
Thr Ala Leu Asn Phe Leu His Leu Tyr His Thr Ile Ile Leu Cys His			
195	200	205	
Thr Ser Glu Arg Lys Glu Ile Leu Ser Leu Asp Lys Leu Glu Ala Gln			
210	215	220	
Leu Lys Ala Cys Asn Cys Arg Leu Ile Phe Ser Lys Ala Lys Pro Ser			
225	230	235	240
Val Leu Ala Leu Cys Leu Leu Asn Leu Glu Val Glu Thr Leu Lys Ser			
245	250	255	
Val Glu Leu Leu Glu Ile Leu Leu Val Lys Lys His Ser Lys Ile			
260	265	270	
Asn Asp Thr Glu Phe Phe Tyr Trp Arg Glu Leu Val Ser Lys Cys Leu			
275	280	285	
Ala Glu Tyr Ser Ser Pro Glu Cys Cys Lys Pro Asp Leu Lys Lys Leu			
290	295	300	

Val Trp Ile Val Ser Arg Arg Thr Ala Gln Asn Leu His Asn Ser Tyr
305 310 315 320

Tyr Ser Val Pro Glu Leu Pro Thr Ile Pro Glu Gly Gly Cys Phe Asp
325 330 335

Glu Ser Glu Ser Glu Asp Ser Cys Glu Asp Met Ser Cys Gly Glu Glu
340 345 350

Ser Leu Ser Ser Ser Pro Pro Ser Asp Gln Glu Cys Thr Phe Phe Phe
355 360 365

Asn Phe Lys Val Ala Gln Thr Leu Cys Phe Pro Ser
370 375 380

<210> 118
<211> 227
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (11)..(11)
<223> X=any amino acid

<400> 118

Met Leu Leu Glu Arg Xaa Gln Cys Asp Gly Xaa Arg Arg Gly Arg Gly
1 5 10 15

Thr Ala Ser Asp Ile Lys Arg Met Glu Lys Ile Ile Ser Glu Lys Leu
20 25 30

His Tyr Glu Leu Glu Ala Thr Thr Ala Leu Asn Phe Leu His Leu Tyr
35 40 45

His Thr Ile Ile Leu Cys His Thr Ser Glu Arg Lys Glu Ile Leu Ser
50 55 60

Leu Asp Lys Leu Glu Ala Gln Leu Lys Ala Cys Asn Cys Arg Leu Ile
65 70 75 80

156

Phe	Ser	Lys	Ala	Lys	Pro	Ser	Val	Leu	Ala	Leu	Cys	Leu	Leu	Asn	Leu
					85						90				95

Glu	Val	Glu	Thr	Leu	Lys	Ser	Val	Glu	Leu	Leu	Glu	Ile	Leu	Leu	Leu
					100						105				110

Val	Lys	Lys	His	Ser	Lys	Ile	Asn	Asp	Thr	Glu	Phe	Phe	Tyr	Trp	Arg
					115					120				125	

Glu	Leu	Val	Ser	Lys	Cys	Leu	Ala	Glu	Tyr	Ser	Ser	Pro	Glu	Cys	Cys
					130				135				140		

Lys	Pro	Asp	Leu	Lys	Lys	Leu	Val	Trp	Ile	Val	Ser	Arg	Arg	Thr	Ala
					145				150			155			160

Gln	Asn	Leu	His	Asn	Ser	Tyr	Tyr	Ser	Val	Pro	Glu	Leu	Pro	Thr	Ile
					165				170			175			

Pro	Glu	Gly	Gly	Cys	Phe	Asp	Glu	Ser	Glu	Ser	Glu	Asp	Ser	Cys	Glu
					180				185			190			

Asp	Met	Ser	Cys	Gly	Glu	Glu	Ser	Leu	Ser	Ser	Ser	Pro	Pro	Ser	Asp
					195				200			205			

Gln	Glu	Cys	Thr	Phe	Phe	Phe	Asn	Phe	Lys	Val	Ala	Gln	Thr	Leu	Cys
					210				215			220			

Phe	Pro	Ser													
225															

.

<210>	119														
<211>	227														
<212>	PRT														
<213>	Homo sapien														

<400>	119														
-------	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Met	Leu	Glu	Arg	Arg	Gln	Cys	Asp	Gly	Leu	Arg	Arg	Gly	Arg	Gly	
1					5				10			15			

Thr	Ala	Ser	Asp	Ile	Lys	Arg	Met	Glu	Lys	Ile	Ile	Ser	Glu	Lys	Leu
					20				25			30			

His	Tyr	Glu	Leu	Glu	Ala	Thr	Thr	Ala	Leu	Asn	Phe	Leu	His	Leu	Tyr
					35				40			45			

His	Thr	Ile	Ile	Leu	Cys	His	Thr	Ser	Glu	Arg	Lys	Glu	Ile	Leu	Ser
					50				55			60			

Leu Asp Lys Leu Glu Ala Gln Leu Lys Ala Cys Asn Cys Arg Leu Ile
65 70 75 80

Phe Ser Lys Ala Lys Pro Ser Val Leu Ala Leu Cys Leu Leu Asn Leu
85 90 95

Glu Val Glu Thr Leu Lys Ser Val Glu Leu Leu Glu Ile Leu Leu Leu
100 105 110

Val Lys Lys His Ser Lys Ile Asn Asp Thr Glu Phe Phe Tyr Trp Arg
115 120 125

Glu Leu Val Ser Lys Cys Leu Ala Glu Tyr Ser Ser Pro Glu Cys Cys
130 135 140

Lys Pro Asp Leu Lys Lys Leu Val Trp Ile Val Ser Arg Arg Thr Ala
145 150 155 160

Gln Asn Leu His Asn Ser Tyr Tyr Ser Val Pro Glu Leu Pro Thr Ile
165 170 175

Pro Glu Gly Gly Cys Phe Asp Glu Ser Glu Ser Glu Asp Ser Cys Glu
180 185 190

Asp Met Ser Cys Gly Glu Glu Ser Leu Ser Ser Ser Pro Pro Ser Asp
195 200 205

Gln Glu Cys Thr Phe Phe Asn Phe Lys Val Ala Gln Thr Leu Cys
210 215 220

Phe Pro Ser
225

<210> 120
<211> 101
<212> PRT
<213> Homo sapien

<400> 120

Met Cys Cys Trp Gln Ala Thr Phe Phe Lys Ala Leu Ser Glu Thr Leu
1 5 10 15

Ile Phe Gly Val Ser Phe Gln Glu Thr Phe Leu Trp Arg Glu Asn Glu
20 25 30

158

Tyr Glu Asp Asn Phe Gln Leu Ile Ile Trp Val Thr Gln Asn Arg Val
 35 40 45

Tyr Gly Tyr Arg Ile Asp Phe Leu Ile Met Ala Ser Asp Val Ala Leu
 50 55 60

Gly Lys Gly Ala Leu Cys Thr Val Cys Ala Cys Met Cys Val Tyr Leu
 65 70 75 80

Tyr Lys Phe Val Ser Phe Gly Met Thr Val Cys Leu Ser Arg Lys Pro
85 90 95

Ile Asn Ser Lys Phe
100

<210> 121
<211> 392
<212> PRT
<213> Homo sapien

<400> 121

Arg Leu Ala Leu Ala Leu Cys Pro Gln Leu Ile Leu Pro His Val Asp
1 5 10 15

Ile Gln Leu Lys Tyr Phe Asp Leu Gly Leu Pro Asn Arg Asp Gln Thr
20 25 30

Asp Asp Gln Val Thr Ile Asp Ser Ala Leu Ala Thr Gln Lys Tyr Ser
35 40 45

Val Ala Val Lys Cys Ala Thr Ile Thr Pro Asp Glu Ala Arg Val Glu
50 55 60

Glu Phe Lys Leu Lys Lys Met Trp Lys Ser Pro Asn Gly Thr Ile Arg
65 70 75 80

Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile Ile Cys Lys Asn
85 90 95

Ile Pro Arg Leu Val Pro Gly Trp Thr Lys Pro Ile Thr Ile Gly Arg
100 105 110

His Ala His Gly Asp Gln Tyr Lys Ala Thr Asp Phe Val Ala Asp Arg
115 120 125

Ala Gly Thr Phe Lys Met Val Phe Thr Pro Lys Asp Gly Ser Gly Val
130 135 140

159

Lys Glu Trp Glu Val Tyr Asn Phe Pro Ala Gly Gly Val Gly Met Gly
145 150 155 160

Met Tyr Asn Thr Asp Glu Ser Ile Ser Gly Phe Ala His Ser Cys Phe
165 170 175

Gln Tyr Ala Ile Gln Lys Trp Pro Leu Tyr Met Ser Thr Lys Asn
180 185 190

Thr Ile Leu Lys Ala Tyr Asp Gly Arg Phe Lys Asp Ile Phe Gln Glu
195 200 205

Ile Phe Asp Lys His Tyr Lys Thr Asp Phe Asp Lys Asn Lys Ile Trp
210 215 220

Tyr Glu His Arg Leu Ile Asp Asp Met Val Ala Gln Val Leu Lys Ser
225 230 235 240

Ser Gly Gly Phe Val Trp Ala Cys Lys Asn Tyr Asp Gly Asp Val Gln
245 250 255

Ser Asp Ile Leu Ala Gln Gly Phe Gly Ser Leu Gly Leu Met Thr Ser
260 265 270

Val Leu Val Cys Pro Asp Gly Lys Thr Ile Glu Ala Glu Ala Ala His
275 280 285

Gly Thr Val Thr Arg His Tyr Arg Glu His Gln Lys Gly Arg Pro Thr
290 295 300

Ser Thr Asn Pro Ile Ala Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu
305 310 315 320

His Arg Gly Lys Leu Asp Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln
325 330 335

Met Leu Glu Lys Val Cys Val Glu Thr Val Glu Ser Gly Ala Met Thr
340 345 350

Lys Asp Leu Ala Gly Cys Ile His Gly Leu Ser Asn Val Lys Leu Asn
355 360 365

Glu His Phe Leu Asn Thr Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn
370 375 380

160

Leu Asp Arg Ala Leu Gly Arg Gln
385 390

<210> 122
<211> 438
<212> PRT
<213> Homo sapien

<400> 122

Met Ala Cys Arg Leu Leu Ile Leu Pro Phe Val Val Met Ser Leu Ser
1 5 10 15

His Trp Gly Asp Ala Leu Leu Leu Ala Leu Cys Pro Gln Leu Ile Leu
20 25 30

Pro His Val Asp Ile Gln Leu Lys Tyr Phe Asp Leu Gly Leu Pro Asn
35 40 45

Arg Asp Gln Thr Asp Asp Gln Val Thr Ile Asp Ser Ala Leu Ala Thr
50 55 60

Gln Lys Tyr Ser Val Ala Val Lys Cys Ala Thr Ile Thr Pro Asp Glu
65 70 75 80

Ala Arg Val Glu Glu Phe Lys Leu Lys Lys Met Trp Lys Ser Pro Asn
85 90 95

Gly Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile
100 105 110

Ile Cys Lys Asn Ile Pro Arg Leu Val Pro Gly Trp Thr Lys Pro Ile
115 120 125

Thr Ile Gly Arg His Ala His Gly Asp Gln Tyr Lys Ala Thr Asp Phe
130 135 140

Val Ala Asp Arg Ala Gly Thr Phe Lys Met Val Phe Thr Pro Lys Asp
145 150 155 160

Gly Ser Gly Val Lys Glu Trp Glu Val Tyr Asn Phe Pro Ala Gly Gly
165 170 175

Val Gly Met Gly Met Tyr Asn Thr Asp Glu Ser Ile Ser Gly Phe Ala
180 185 190

His Ser Cys Phe Gln Tyr Ala Ile Gln Lys Lys Trp Pro Leu Tyr Met

161

195

200

205

Ser Thr Lys Asn Thr Ile Leu Lys Ala Tyr Asp Gly Arg Phe Lys Asp
210 215 220

Ile Phe Gln Glu Ile Phe Asp Lys His Tyr Lys Thr Asp Phe Asp Lys
225 230 235 240

Asn Lys Ile Trp Tyr Glu His Arg Leu Ile Asp Asp Met Val Ala Gln
245 250 255

Val Leu Lys Ser Ser Gly Gly Phe Val Trp Ala Cys Lys Asn Tyr Asp
260 265 270

Gly Asp Val Gln Ser Asp Ile Leu Ala Gln Gly Phe Gly Ser Leu Gly
275 280 285

Leu Met Thr Ser Val Leu Val Cys Pro Asp Gly Lys Thr Ile Glu Ala
290 295 300

Glu Ala Ala His Gly Thr Val Thr Arg His Tyr Arg Glu His Gln Lys
305 310 315 320

Gly Arg Pro Thr Ser Thr Asn Pro Ile Ala Ser Ile Phe Ala Trp Thr
325 330 335

Arg Gly Leu Glu His Arg Gly Lys Leu Asp Gly Asn Gln Asp Leu Ile
340 345 350

Arg Phe Ala Gln Met Leu Glu Lys Val Cys Val Glu Thr Val Glu Ser
355 360 365

Gly Ala Met Thr Lys Asp Leu Ala Gly Cys Ile His Gly Leu Ser Asn
370 375 380

Val Lys Leu Asn Glu His Phe Leu Asn Thr Thr Asp Phe Leu Asp Thr
385 390 395 400

Ile Lys Ser Asn Leu Asp Ser Ser Pro Gly Gln Ala Val Gly Gly
405 410 415

Ala Thr His Gly Cys Ser Gly Gly Ala Arg Ala Glu Pro Ala Gly Pro
420 425 430

Pro Glu Arg Gly Arg Gly
435

<210> 123
<211> 292
<212> PRT
<213> Homo sapien

<400> 123

Pro Gly His Pro Pro Thr Gly Ala Pro Arg Leu Ala Ile Leu Leu Ser
1 5 10 15

Leu Gln Tyr Lys Ala Thr Asp Phe Val Ala Asp Arg Ala Gly Thr Phe
20 25 30

Lys Met Val Phe Thr Pro Lys Asp Gly Ser Gly Val Lys Glu Trp Glu
35 40 45

Val Tyr Asn Phe Pro Ala Gly Gly Val Gly Met Gly Met Tyr Asn Thr
50 55 60

Asp Glu Ser Ile Ser Gly Phe Ala His Ser Cys Phe Gln Tyr Ala Ile
65 70 75 80

Gln Lys Lys Trp Pro Leu Tyr Met Ser Thr Lys Asn Thr Ile Leu Lys
85 90 95

Ala Tyr Asp Gly Arg Phe Lys Asp Ile Phe Gln Glu Ile Phe Asp Lys
100 105 110

His Tyr Lys Thr Asp Phe Asp Lys Asn Lys Ile Trp Tyr Glu His Arg
115 120 125

Leu Ile Asp Asp Met Val Ala Gln Val Leu Lys Ser Ser Gly Gly Phe
130 135 140

Val Trp Ala Cys Lys Asn Tyr Asp Gly Asp Val Gln Ser Asp Ile Leu
145 150 155 160

Ala Gln Gly Phe Gly Ser Leu Gly Leu Met Thr Ser Val Leu Val Cys
165 170 175

Pro Asp Gly Lys Thr Ile Glu Ala Glu Ala Ala His Gly Thr Val Thr
180 185 190

Arg His Tyr Arg Glu His Gln Lys Gly Arg Pro Thr Ser Thr Asn Pro
195 200 205

163

Ile Ala Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu His Arg Gly Lys
210 215 220

Leu Asp Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln Met Leu Glu Lys
 225 230 235 240

Val Cys Val Glu Thr Val Glu Ser Gly Ala Met Thr Lys Asp Leu Ala
245 250 255

Gly Cys Ile His Gly Leu Ser Asn Val Lys Leu Asn Glu His Phe Leu
260 265 270

Asn Thr Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn Leu Asp Arg Ala
275 280 285

Leu Gly Arg Gln
290

<210> 124
<211> 417
<212> PRT
<213> Homo sapien

<400> 124

Met Lys Asn Phe Arg Thr Pro Val Trp Leu Cys Cys Cys Leu Gly Phe
 1 5 10 15

Lys Phe Trp Leu Lys Asp Gly Gly Cys Ser Gly Thr Thr Ile Ile Ser
20 25 30

Val Leu Thr Glu Phe Lys Leu Lys Lys Met Trp Lys Ser Pro Asn Gly
35 40 45

Thr Ile Arg Asn Ile Leu Gly Gly Thr Val Phe Arg Glu Pro Ile Ile
50 55 60

Cys Lys Asn Ile Pro Arg Leu Val Pro Gly Trp Thr Lys Pro Ile Thr
65 70 75 80

Ile Gly Arg His Ala His Gly Asp Gln Val Gly Gln Gly Gly Glu Gly
85 90 95

Ile His Arg Pro Gly His Pro Pro Thr Gly Ala Pro Arg Leu Ala Ile
100 105 110

Leu Leu Ser Leu Gln Tyr Lys Ala Thr Asp Phe Val Ala Asp Arg Ala
115 120 125

Gly Thr Phe Lys Met Val Phe Thr Pro Lys Asp Gly Ser Gly Val Lys
130 135 140

Glu Trp Glu Val Tyr Asn Phe Pro Ala Gly Gly Val Gly Met Gly Met
145 150 155 160

Tyr Asn Thr Asp Glu Ser Ile Ser Gly Phe Ala His Ser Cys Phe Gln
165 170 175

Tyr Ala Ile Gln Lys Lys Trp Pro Leu Tyr Met Ser Thr Lys Asn Thr
180 185 190

Ile Leu Lys Ala Tyr Asp Gly Arg Phe Lys Asp Ile Phe Gln Glu Ile
195 200 205

Phe Asp Lys His Tyr Lys Thr Asp Phe Asp Lys Asn Lys Ile Trp Tyr
210 215 220

Glu His Arg Leu Ile Asp Asp Met Val Ala Gln Val Leu Lys Ser Ser
225 230 235 240

Gly Gly Phe Val Trp Ala Cys Lys Asn Tyr Asp Gly Asp Val Gln Ser
245 250 255

Asp Ile Leu Ala Gln Gly Phe Gly Ser Leu Gly Leu Met Thr Ser Val
260 265 270

Leu Val Cys Pro Asp Gly Lys Thr Ile Glu Ala Glu Ala Ala His Gly
275 280 285

Thr Val Thr Arg His Tyr Arg Glu His Gln Lys Gly Arg Pro Thr Ser
290 295 300

Thr Asn Pro Ile Ala Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu His
305 310 315 320

Arg Gly Lys Leu Asp Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln Met
325 330 335

Leu Glu Lys Val Cys Val Glu Thr Val Glu Ser Gly Ala Met Thr Lys
340 345 350

Asp Leu Ala Gly Cys Ile His Gly Leu Ser Asn Val Lys Leu Asn Glu
355 360 365

165

His Phe Leu Asn Thr Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn Leu
370 375 380

Asp Ser Ser Pro Gly Gln Ala Val Gly Gly Ala Thr His Gly Cys
385 390 395 400

Ser Gly Gly Ala Arg Ala Glu Pro Ala Gly Pro Pro Glu Arg Gly Arg
405 410 415

Gly

<210> 125

<211> 255

<212> PRT

<213> Homo sapien

<400> 125

Lys Pro Thr Met Gly Val Ser Arg Thr Ser Ser Arg Arg Ser Leu Thr
1 5 10 15

Ser Lys Ala Ser Ser Met Tyr Ser Val Ala Phe Leu Pro Phe Pro Pro
20 25 30

Cys Cys Ser His Pro Thr Leu Gly Arg Ser Leu Leu Glu Cys Ile Trp
35 40 45

Leu Ser Ser Glu Ala Gln Gly Gly Ile Pro Asn Leu Ser Ala Phe Cys
50 55 60

Pro Leu Pro Ile Thr Asp Leu Phe Thr Pro Arg His Tyr Lys Thr Asp
65 70 75 80

Phe Asp Lys Asn Lys Ile Trp Tyr Glu His Arg Leu Ile Asp Asp Met
85 90 95

Val Ala Gln Val Leu Lys Ser Ser Gly Gly Phe Val Trp Ala Cys Lys
100 105 110

Asn Tyr Asp Gly Asp Val Gln Ser Asp Ile Leu Ala Gln Gly Phe Gly
115 120 125

Ser Leu Gly Leu Met Thr Ser Val Leu Val Cys Pro Asp Gly Lys Thr
130 135 140

Ile Glu Ala Glu Ala Ala His Gly Thr Val Thr Arg His Tyr Arg Glu

166

145

150

155

160

His Gln Lys Gly Arg Pro Thr Ser Thr Asn Pro Ile Ala Ser Ile Phe
165 170 175

Ala Trp Thr Arg Gly Leu Glu His Arg Gly Lys Leu Asp Gly Asn Gln
180 185 190

Asp Leu Ile Arg Phe Ala Gln Met Leu Glu Lys Val Cys Val Glu Thr
195 200 205

Val Glu Ser Gly Ala Met Thr Lys Asp Leu Ala Gly Cys Ile His Gly
210 215 220

Leu Ser Asn Val Lys Leu Asn Glu His Phe Leu Asn Thr Thr Asp Phe
225 230 235 240

Leu Asp Thr Ile Lys Ser Asn Leu Asp Arg Ala Leu Gly Arg Gln
245 250 255

<210> 126

<211> 289

<212> PRT

<213> Homo sapien

<400> 126

Met Ser Thr Lys Asn Thr Ile Leu Lys Ala Tyr Asp Gly Arg Phe Lys
1 5 10 15

Asp Ile Phe Gln Glu Ile Phe Asp Asn Lys Ala Ser Ser Met Tyr Ser
20 25 30

Val Ala Phe Leu Pro Phe Pro Pro Cys Cys Ser His Pro Thr Leu Gly
35 40 45

Arg Ser Leu Leu Glu Cys Ile Trp Leu Ser Ser Glu Ala Gln Gly Gly
50 55 60

Ile Pro Asn Leu Ser Ala Phe Cys Pro Leu Pro Ile Thr Asp Leu Phe
65 70 75 80

Thr Pro Arg His Tyr Lys Thr Asp Phe Asp Lys Asn Lys Ile Trp Tyr
85 90 95

Glu His Arg Leu Ile Asp Asp Met Val Ala Gln Val Leu Lys Ser Ser
100 105 110

167

Gly Gly Phe Val Trp Ala Cys Lys Asn Tyr Asp Gly Asp Val Gln Ser
115 120 125

Asp Ile Leu Ala Gln Gly Phe Gly Ser Leu Gly Leu Met Thr Ser Val
130 135 140

Leu Val Cys Pro Asp Gly Lys Thr Ile Glu Ala Glu Ala Ala His Gly
145 150 155 160

Thr Val Thr Arg His Tyr Arg Glu His Gln Lys Gly Arg Pro Thr Ser
165 170 175

Thr Asn Pro Ile Ala Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu His
 180 185 186

Arg Gly Lys Leu Asp Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln Met
195 200 205

Leu Glu Lys Val Cys Val Glu Thr Val Glu Ser Gly Ala Met Thr Lys
210 215

Asp Leu Ala Gly Cys Ile His Gly Leu Ser Asn Val Lys Leu Asn Glu
225 230 235

His Phe Leu Asn Thr Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn Leu
245 256

Asp Ser Ser Pro Gly Gln Ala Val Gly Gly Gly Ala Thr His Gly Cys
260 265

Ser Gly Gly Ala Arg Ala Glu Pro Ala Gly Pro Pro Glu Arg Arg Gly Arg
³⁷⁵

二

<210> 127
<211> 167
<212> PRT
<213> *Homo sapien*

<400> 127

Val Glu Pro Arg Thr Met Ala Ala Thr Ile Leu Gly Cys Arg Gly Gln
1 5 10 15

Gln Gly Ser Ala Gly Trp Pro Gln Glu Arg Arg Gly Pro Glu Arg Lys

168

20

25

30

Ala Phe Tyr Pro Pro Gly Phe Gly Ser Leu Gly Leu Met Thr Ser Val
35 40 45

Leu Val Cys Pro Asp Gly Lys Thr Ile Glu Ala Glu Ala Ala His Gly
50 55 60

Thr Val Thr Arg His Tyr Arg Glu His Gln Lys Gly Arg Pro Thr Ser
65 70 75 80

Thr Asn Pro Ile Ala Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu His
85 90 95

Arg Gly Lys Leu Asp Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln Met
100 105 110

Leu Glu Lys Val Cys Val Glu Thr Val Glu Ser Gly Ala Met Thr Lys
115 120 125

Asp Leu Ala Gly Cys Ile His Gly Leu Ser Asn Val Lys Leu Asn Glu
130 135 140

His Phe Leu Asn Thr Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn Leu
145 150 155 160

Asp Arg Ala Leu Gly Arg Gln
165

<210> 128
<211> 188
<212> PRT
<213> Homo sapien

<400> 128

Met Ala Ala Thr Ile Leu Gly Cys Arg Gly Gln Gln Gly Ser Ala Gly
1 5 10 15

Trp Pro Gln Glu Arg Arg Gly Pro Glu Arg Lys Ala Phe Tyr Pro Pro
20 25 30

Gly Phe Gly Ser Leu Gly Leu Met Thr Ser Val Leu Val Cys Pro Asp
35 40 45

Gly Lys Thr Ile Glu Ala Glu Ala Ala His Gly Thr Val Thr Arg His
50 55 60

169

Tyr Arg Glu His Gln Lys Gly Arg Pro Thr Ser Thr Asn Pro Ile Ala
65 70 75 80

Ser Ile Phe Ala Trp Thr Arg Gly Leu Glu His Arg Gly Lys Leu Asp
85 90 95

Gly Asn Gln Asp Leu Ile Arg Phe Ala Gln Met Leu Glu Lys Val Cys
100 105 110

Val Glu Thr Val Glu Ser Gly Ala Met Thr Lys Asp Leu Ala Gly Cys
115 120 125

Ile His Gly Leu Ser Asn Val Lys Leu Asn Glu His Phe Leu Asn Thr
130 135 140

Thr Asp Phe Leu Asp Thr Ile Lys Ser Asn Leu Asp Ser Ser Pro Gly
145 150 155 160

Gln Ala Val Gly Gly Ala Thr His Gly Cys Ser Gly Gly Ala Arg
165 170 175

Ala Glu Pro Ala Gly Pro Pro Glu Arg Gly Arg Gly
180 185

<210> 129
<211> 162
<212> PRT
<213> Homo sapien

<400> 129

Pro Ala Arg Pro Ala Pro Ala Arg Pro Ser Val Ser Val Ser Pro Arg
1 5 10 15

Pro Gly Ser Arg Glu Glu Arg Arg Ala Leu Gly Pro Leu Pro Pro Cys
20 25 30

Ser Phe Ala Leu Gln Leu Gly Met Ala Gly Tyr Leu Arg Val Val Arg
35 40 45

Ser Leu Cys Arg Ala Ser Gly Ser Arg Pro Ala Trp Ala Pro Ala Ala
50 55 60

Leu Thr Ala Pro Thr Ser Gln Glu Gln Pro Arg Arg His Tyr Ala Asp
65 70 75 80

Lys Arg Ile Lys Val Ala Lys Pro Val Val Glu Met Asp Gly Asp Glu

85	90	95
	170	
Met Thr Arg Ile Ile Trp Gln Phe Ile Lys Glu Lys Cys Glu Ala Glu		
100	105	110

115	120	125
Arg Ala Leu Pro Glu His His Gly Leu Pro Arg His His Gln Glu Gln		

130	135	140
Pro Gly Gln Ser Pro Gly Gln Ala Val Gly Gly Ala Thr His Gly		

145	150	155
Cys Ser Gly Gly Ala Arg Ala Glu Pro Ala Gly Pro Pro Glu Arg Gly		
		160

Arg Gly

<210> 130
<211> 112
<212> PRT
<213> Homo sapien

<400> 130

1	5	10
Met Ala Gly Tyr Leu Arg Val Val Arg Ser Leu Cys Arg Ala Ser Gly		
		15

20	25	30
Ser Arg Pro Ala Trp Ala Pro Ala Ala Leu Thr Ala Pro Thr Ser Gln		

35	40	45
Glu Gln Pro Arg Arg His Tyr Ala Asp Lys Arg Ile Lys Val Ala Lys		

50	55	60
Pro Val Val Glu Met Asp Gly Asp Glu Met Thr Arg Ile Ile Trp Gln		

65	70	75
Phe Ile Lys Glu Lys Cys Glu Ala Glu Arg Ala Leu Pro Glu His His		
		80

85	90	95
Gly Leu Pro Arg His His Gln Glu Gln Pro Gly Gln Gln Pro Trp Ala		

100	105	110
Gly Ser Arg Gly Arg Arg His Pro Trp Leu Gln Trp Arg Gly Gln Gly		

<210> 131
<211> 306
<212> PRT

171

<213> Homo sapien

<400> 131

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
20 25 30

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
65 70 75 80

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
130 135 140

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
165 170 175

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe

172

225 230 235 240

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln
275 280 285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
290 295 300

Ser Ser
305

<210> 132
<211> 508
<212> PRT
<213> Homo sapien

<400> 132

Met Pro Trp Arg Ala Pro Ser Ala Ser Ser Ala Ser Ala Gly Arg Ile
1 5 10 15

Leu Leu Arg Pro Thr Glu Glu Glu Gly Gly Ala Glu Arg Ser Phe Ser
20 25 30

Gly Pro Arg Gly Ser Ser Gly Arg Ile Pro Arg Phe Val Ser Ile Ser
35 40 45

Ile Thr Asn Gly Pro Val Phe Cys Gly Val Val Gly Ala Val Ala Arg
50 55 60

His Glu Tyr Thr Val Ile Gly Pro Lys Val Ser Leu Ala Ala Arg Met
65 70 75 80

Ile Thr Ala Tyr Pro Gly Leu Val Ser Cys Asp Glu Val Thr Tyr Leu
85 90 95

Arg Ser Met Leu Pro Ala Tyr Asn Phe Lys Lys Leu Pro Glu Lys Met
100 105 110

Met Lys Asn Ile Ser Asn Pro Gly Lys Ile Tyr Glu Tyr Leu Gly His
115 120 125

173

Arg Arg Cys Ile Met Phe Gly Lys Arg His Leu Ala Arg Lys Arg Asn
130 135 140

Lys Asn His Pro Leu Leu Gly Val Leu Gly Ala Pro Cys Leu Ser Thr
145 150 155 160

Asp Trp Glu Lys Glu Leu Glu Ala Phe Gln Met Ala Gln Gln Gly Cys
165 170 175

Leu His Gln Lys Lys Gly Gln Ala Val Leu Tyr Glu Gly Gly Lys Gly
180 185 190

Tyr Gly Lys Ser Gln Leu Leu Ala Glu Ile Asn Phe Leu Ala Gln Lys
195 200 205

Glu Gly His Ser Tyr Pro Ser Gln Val Leu Trp Lys Pro Thr Leu Phe
210 215 220

Glu Val Leu Cys Gln Asp Leu Leu Ser Lys Asp Val Leu Leu Phe His
 225 230 235 240

Val Leu Gln Lys Glu Glu Glu Asn Ser Lys Trp Glu Thr Leu Ser
245 250 255

Ala Asn Ala Met Lys Ser Ile Met Tyr Ser Ile Ser Pro Ala Asn Ser
260 265 270

Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr Val Lys Asp Asp Val Asn
275 280

Leu Asp Thr Val Leu Leu Leu Pro Phe Leu Lys Glu Ile Ala Val Ser
280 281

Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln Leu Leu Val Lys Cys Ala
205

Ala Ile Ile Gly His Ser Phe His Ile Asp Leu Leu Gln His Leu Leu

Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln Val Leu Arg Ala Leu Val

Asp Ile His Val Leu Cys Trp Ser Asp Lys Ser Glu Glu Ile

174

Glu Pro Ile Leu Met Pro Ser Ser Ile Asp Ile Ile Asp Gly Thr Lys
370 375 380

Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser Ala Ser Leu Leu Arg Leu
385 390 395 400

Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu Val Leu Glu Phe Gly Val
405 410 415

Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu Trp Pro Lys Glu Gln Gln
420 425 430

Ile Ala Leu His Leu Glu Cys Ala Cys Phe Leu Gln Val Leu Ala Cys
435 440 445

Arg Cys Gly Ser Cys His Gly Gly Asp Phe Val Pro Phe His His Phe
450 455 460

Ala Val Cys Ser Thr Lys Asn Ser Lys Gly Thr Ser Arg Phe Cys Thr
465 470 475 480

Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln Val Ile Thr Glu Lys Leu
485 490 495

Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys Ser Ser
500 505

<210> 133
<211> 306
<212> PRT
<213> Homo sapien

<400> 133

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
20 25 30

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
65 70 75 80

175

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
130 135 140

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
165 170 175

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe
225 230 235 240

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln
275 280 285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
290 295 300

Ser Ser
305

176

<210> 134
<211> 429
<212> PRT
<213> Homo sapien

<400> 134

Met Ile Thr Ala Tyr Pro Gly Leu Val Ser Cys Asp Glu Val Thr Tyr
1 5 10 15

Leu Arg Ser Met Leu Pro Ala Tyr Asn Phe Lys Lys Leu Pro Glu Lys
20 25 30

Met Met Lys Asn Ile Ser Asn Pro Gly Lys Ile Tyr Glu Tyr Leu Gly
35 40 45

His Arg Arg Cys Ile Met Phe Gly Lys Arg His Leu Ala Arg Lys Arg
50 55 60

Asn Lys Asn His Pro Leu Leu Gly Val Leu Gly Ala Pro Cys Leu Ser
65 70 75 80

Thr Asp Trp Glu Lys Glu Leu Glu Ala Phe Gln Met Ala Gln Gln Gly
85 90 95

Cys Leu His Gln Lys Lys Gly Gln Ala Val Leu Tyr Glu Gly Gly Lys
100 105 110

Gly Tyr Gly Lys Ser Gln Leu Leu Ala Glu Ile Asn Phe Leu Ala Gln
115 120 125

Lys Glu Gly His Ser Tyr Pro Ser Gln Val Leu Trp Lys Pro Thr Leu
130 135 140

Phe Glu Val Leu Cys Gln Asp Leu Leu Ser Lys Asp Val Leu Leu Phe
145 150 155 160

His Val Leu Gln Lys Glu Glu Glu Asn Ser Lys Trp Glu Thr Leu
165 170 175

Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser Ile Ser Pro Ala Asn
180 185 190

Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr Val Lys Asp Asp Val
195 200 205

Asn Leu Asp Thr Val Leu Leu Pro Phe Leu Lys Glu Ile Ala Val

177

210

215

220

Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln Leu Leu Val Lys Cys
 225 230 235 240

Ala Ala Ile Ile Gly His Ser Phe His Ile Asp Leu Leu Gln His Leu
 245 250 255

Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln Val Leu Arg Ala Leu
 260 265 270

Val Asp Ile His Val Leu Cys Trp Ser Asp Lys Ser Gln Glu Leu Pro
 275 280 285

Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp Ile Ile Asp Gly Thr
 290 295 300

Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser Ala Ser Leu Leu Arg
 305 310 315 320

Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu Val Leu Glu Phe Gly
 325 330 335

Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu Trp Pro Lys Glu Gln
 340 345 350

Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe Leu Gln Val Leu Ala
 355 360 365

Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe Val Pro Phe His His
 370 375 380

Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly Thr Ser Arg Phe Cys
 385 390 395 400

Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln Val Ile Thr Glu Lys
 405 410 415

Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys Ser Ser
 420 425

<210> 135
<211> 306
<212> PRT
<213> Homo sapien

<400> 135

178

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
20 25 30

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
65 70 75 80

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
130 135 140

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
165 170 175

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe
225 230 235 240

179

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln
275 280 285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
290 295 300

Ser Ser
305

<210> 136
<211> 306
<212> PRT
<213> Homo sapien

<400> 136

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
20 25 30

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
65 70 75 80

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
130 135 140

180

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
165 170 175

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe
225 230 235 240

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln
275 280 285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
290 295 300

Ser Ser
305

<210> 137
<211> 306
<212> PRT
<213> Homo sapien

<400> 137

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
20 25 30

181

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
65 70 75 80

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
130 135 140

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
165 170 175

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe
225 230 235 240

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln

182

275

280

285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
 290 295 300

Ser Ser
 305

<210> 138
 <211> 306
 <212> PRT
 <213> Homo sapien

<400> 138

Thr Phe Trp His Arg Lys Lys Gly Ile Ala Thr Leu His Arg Cys Phe
 1 5 10 15

Gly Asn Pro Leu Tyr Cys Glu Val Leu Cys Gln Asp Leu Leu Ser Lys
 20 25 30

Asp Val Leu Leu Phe His Val Leu Gln Lys Glu Glu Glu Asn Ser
 35 40 45

Lys Trp Glu Thr Leu Ser Ala Asn Ala Met Lys Ser Ile Met Tyr Ser
 50 55 60

Ile Ser Pro Ala Asn Ser Glu Glu Gly Gln Glu Leu Tyr Val Cys Thr
 65 70 75 80

Val Lys Asp Asp Val Asn Leu Asp Thr Val Leu Leu Leu Pro Phe Leu
 85 90 95

Lys Glu Ile Ala Val Ser Gln Leu Asp Gln Leu Ser Pro Glu Glu Gln
 100 105 110

Leu Leu Val Lys Cys Ala Ala Ile Ile Gly His Ser Phe His Ile Asp
 115 120 125

Leu Leu Gln His Leu Leu Pro Gly Trp Asp Lys Asn Lys Leu Leu Gln
 130 135 140

Val Leu Arg Ala Leu Val Asp Ile His Val Leu Cys Trp Ser Asp Lys
 145 150 155 160

Ser Gln Glu Leu Pro Ala Glu Pro Ile Leu Met Pro Ser Ser Ile Asp
 165 170 175

183

Ile Ile Asp Gly Thr Lys Glu Lys Lys Thr Lys Leu Asp Gly Gly Ser
180 185 190

Ala Ser Leu Leu Arg Leu Gln Glu Glu Leu Ser Leu Pro Gln Thr Glu
195 200 205

Val Leu Glu Phe Gly Val Pro Leu Leu Arg Ala Ala Ala Trp Glu Leu
210 215 220

Trp Pro Lys Glu Gln Gln Ile Ala Leu His Leu Glu Cys Ala Cys Phe
225 230 235 240

Leu Gln Val Leu Ala Cys Arg Cys Gly Ser Cys His Gly Gly Asp Phe
245 250 255

Val Pro Phe His His Phe Ala Val Cys Ser Thr Lys Asn Ser Lys Gly
260 265 270

Thr Ser Arg Phe Cys Thr Tyr Arg Asp Thr Gly Ser Val Leu Thr Gln
275 280 285

Val Ile Thr Glu Lys Leu Gln Leu Pro Ser Pro Gln Glu Gln Arg Lys
290 295 300

Ser Ser
305

<210> 139
<211> 121
<212> PRT
<213> Homo sapien

<400> 139

Met Arg Ser Thr Arg Glu Arg Arg Pro Gln Glu Arg Arg Arg Gln Gly
1 5 10 15

Ser Val Arg Gln Gly Arg Thr Gly Gly Ser Arg Phe Ala Ile Ile Pro
20 25 30

Gly Ser Arg Leu Cys Phe Val Gly Pro Ser His Cys Ile Leu Ala His
35 40 45

Thr Gly Glu Phe Trp Pro Trp Glu Asn Trp Ser Gln His Ala Ala Lys
50 55 60

Leu Ser His Gly Arg Gln Arg Ile Pro Thr His Cys Arg Ser Lys Pro

184

65 70 75 80
Cys Trp Lys Lys Gln Asn Ser Ser Pro Ser Val Glu Leu Arg Gly Asp
85 90 95

Trp Ser Arg Ala Pro Ala Asp Thr Lys Ile Gln Val Ala Gln Val Ser
100 105 110

His Arg Lys Trp Arg Ser Ile Cys Thr
115 120

<210> 140
<211> 125
<212> PRT
<213> Homo sapien

<400> 140

Glu Phe Gly Gly Val Gly Ser Lys Leu Asn Thr Ala Ala Val His Gly
1 5 10 15

Arg Asn Tyr Ser Ile His Thr Phe Ser Glu Tyr Pro Ile Thr Lys Ala
20 25 30

Lys Lys Asn Thr Lys Gly Phe Val Leu Leu Leu Gly Val Asp Leu Ile
35 40 45

Pro Arg Gln Ser Ser Gly His Arg His Arg Gly Cys Ala Gln Ala Cys
50 55 60

Pro Gln Pro Tyr Ala Ala Val Glu Ser Gly Arg Leu Leu Gln Asp Cys
65 70 75 80

Trp Pro Ser Pro Arg Met Ser Ala Ser Phe Ser Ile Tyr Trp Leu Leu
85 90 95

Leu Leu Tyr Val Met Leu Thr Leu Leu Leu Asn Thr Gly Leu Phe Ala
100 105 110

Phe Phe Pro Leu Met Glu Thr Trp Glu Arg His Tyr Phe
115 120 125

<210> 141
<211> 764
<212> PRT
<213> Homo sapien

<400> 141

185

Met Gln Ser Ser Leu Tyr Phe Glu Arg Ile Lys Tyr Asp Leu Gln Lys
1 5 10 15

Leu His Gly Gly Leu Ser Lys Thr Leu Asn Tyr Leu Phe Phe Val Glu
20 25 30

Lys Ser Tyr Phe Arg His His Phe Ile Pro Gln Gln Leu Ala Val Lys
35 40 45

Pro Leu Leu Cys Cys Met Pro Val Thr Leu Leu Asp Cys Gly Asp Tyr
50 55 60

Gln Cys Ser Arg Leu Leu Arg Ala Arg Val Gly Trp Gly Ile Lys Thr
65 70 75 80

Gly Lys Gln Ile Ala Thr Ile Leu Tyr Cys Glu Cys Leu Cys Trp Arg
85 90 95

Lys Tyr Arg Glu Leu Leu Glu His Leu Arg Gly Ala Pro Thr Leu Asn
100 105 110

Leu Gly Val Ser Arg Gly Ile Leu Lys Lys Val Lys Ala Lys Pro Gln
115 120 125

Ser Ile Ser Ser Leu Gly Ile Glu Gln Asn Val Arg Gly Glu Glu Met
130 135 140

Pro Lys Ala Arg Arg Glu Glu Tyr Ser Lys Gln Glu Gly Phe Gln Arg
145 150 155 160

Glu Lys Ser Ile Pro Asn Asn Ile Cys Thr Asn Leu Met Gly Arg Glu
165 170 175

Asn Val Gly Trp Gly Trp Met Met Arg Leu Lys Lys Ala Arg Ser
180 185 190

Glu Ile Ile Ser Gly Leu Val His His Val Lys Glu Cys Arg Leu Asp
195 200 205

Ser Val Val Asn Arg Lys Ala Ala Gln Phe Ile Met Asn Ile Leu Glu
210 215 220

Asp Ser His Trp Asn Met Glu Asn Lys Val Gly Asp Asp Tyr Ile Leu
225 230 235 240

Glu Ala Gly Arg Thr Phe Leu Arg Lys Leu His Tyr Phe Gly Glu Asn

245	250	255
Asp Gly His Lys His Glu Glu Leu Glu Val Ile Met Thr Ser Ser Leu		
260	265	270
Ile Phe Gln Lys Gly Phe Gly Arg Tyr Asn Ile Gly Thr Leu Thr Gly		
275	280	285
Leu Thr Lys Gly Asp Glu Ile His His Ile Asn Cys Gln Thr Gln Gly		
290	295	300
Gln Met Ser Asn Tyr Phe Ala Tyr Asp Val Glu Ile Thr Asn Phe Ser		
305	310	315
320		
Ser Gly Asn Gln Lys Leu Gln Asn Leu Val Phe Pro Ser Pro Arg Ile		
325	330	335
Leu Ser Val Gln Thr Ile Cys Thr Thr Pro Pro Ile Ser Leu Pro Leu		
340	345	350
His Val Cys Pro Thr Ser Lys Ser Arg Ser Ile His Thr Gly Lys Thr		
355	360	365
Arg Ala Val Gln Val Ser Glu Asn Glu Lys Glu Glu Leu Ser Cys Ala		
370	375	380
Glu Pro Ile Gln Asn Lys His Ile Leu Cys Ile Asp Ser Trp Asn Leu		
385	390	395
400		
Glu Arg Asn Ser Pro Asn Ser Ile Gly Ile Trp Met Val Cys Asn Pro		
405	410	415
Trp Leu Gly Ser Ala Phe Lys Lys Pro Tyr Leu Glu Ile Pro Ser Met		
420	425	430
Glu Pro Ser Ser Ile Lys Ala His Leu Lys Ala Tyr Ile Lys Asn Lys		
435	440	445
Ile Leu Ala Ala Leu Tyr Thr Asn Asn Asp Val Met Ile Lys Leu Ser		
450	455	460
Asp Ala Ile Ile Lys Trp Asn Tyr Lys Met Val Tyr Pro Leu Gln Lys		
465	470	475
480		
Lys Lys Ala Lys Phe Ser Val Glu His Cys Asp Phe Met Ser Leu His		
485	490	495

Ser Leu Gly Ala Glu Glu Gly Ala Leu Val Ser Ser Glu Val Glu Glu
500 505 510

Lys Thr Trp Arg Leu Ile Ile Tyr Ala Met Phe Phe His Leu Lys Glu
515 520 525

Ala Phe Phe Leu Asp Tyr Leu Ile Gln Phe Pro Ser Arg Lys Leu Leu
530 535 540

Val Pro Leu Thr Arg Gln Gln Leu Gly Arg Gln Lys Leu Tyr Cys Met
545 550 555 560

Tyr Met Val Ala Val Gly Arg Arg Phe Leu Ser Pro Gly Pro His Trp
565 570 575

Pro Tyr Thr Ser Pro Leu Leu Val Met Pro Gly His Arg Pro Pro Val
580 585 590

Ala Ile Ile Ser Tyr Leu Ser Leu Trp Leu Val Asn Leu Ser Ile Leu
595 600 605

Ser Ala Ser Ala Leu Gln Ser Ala Gly Thr Leu Leu Thr Ser Ile Ser
610 615 620

Cys Trp Leu Ser Thr Phe Leu Ile Gly Pro Ala Leu Phe Ser Ser Gly
625 630 635 640

Pro Ala Val Glu Ser Pro Cys Pro Phe Arg Arg Ala Met Ala Tyr His
645 650 655

Cys Leu Leu Ser Leu His Ser Ala Ala Thr Thr Leu Asn Pro Ser Phe
660 665 670

Ser Lys Asp Val Ala Asp Phe Thr Gly Lys His Lys Arg Leu Asp Leu
675 680 685

Pro Gly Leu Pro Phe Thr Cys Leu Asn Leu Thr Ser Phe Asn Phe Gln
690 695 700

Ser Gln Asn Val Gly Ile Val Ser Ser Leu Pro Tyr Ile Phe Leu Leu
705 710 715 720

Leu Asn His Glu Ser Leu Ser Leu Pro Leu Ala Met Cys Trp Arg Leu
725 730 735

188

Leu Ser Gly Phe Arg Met Ser Ser His Leu Val Leu Val Ala Phe Asp
740 745 750

Ala Ser Ser Pro Pro Phe Lys Asp Thr Phe Glu Ile
755 760

<210> 142
<211> 267
<212> PRT
<213> Homo sapien

<400> 142

Val Arg Ala Pro Ser Pro Gly Gln Ala Gly Arg Ala Glu Gly Ala Asp
1 5 10 15

Pro Gln Pro Gly Pro Ala His Leu His Asp Gly Ser Glu Leu Leu Arg
20 25 30

Gly Lys Leu Arg Gln Leu Ser Glu Asp Asn Val Arg Pro Arg Gly Ala
35 40 45

Arg Leu Ser Ser Gly Pro Gly Thr Gly Val Ser Val Leu Phe Glu Arg
50 55 60

Asp Gly Glu Leu His Phe Pro Ala Cys His Arg Ala Leu Arg Ala Cys
65 70 75 80

Asp Gly Lys Ser Ser Ser Gln Pro Asn Val Ile Ser Ala Ala Leu Leu
85 90 95

Gly Pro Arg Ser Val Val Ser Gly Gly Leu Val Trp Arg Pro Val
100 105 110

Ser Gly Phe Gly Asp Gly Ser Asp Ala Ile Thr Ala Arg Gln Gly Val
115 120 125

Ser Arg Gly Val Lys Ala Ala Met Asn Arg Val Leu Cys Ala Pro Ala
130 135 140

Ala Gly Ala Val Arg Ala Leu Arg Leu Ile Gly Trp Ala Ser Arg Ser
145 150 155 160

Leu His Pro Leu Pro Gly Ser Arg Asp Arg Ala His Pro Ala Ala Glu
165 170 175

Glu Glu Asp Asp Pro Asp Arg Pro Ile Glu Phe Ser Ser Ser Lys Ala

189

180

185

190

Asn Pro His Arg Trp Ser Val Gly His Thr Met Gly Lys Gly His Gln
195 200 205

Arg Pro Trp Trp Lys Val Leu Pro Leu Ser Cys Phe Leu Val Ala Leu
210 215 220

Ile Ile Trp Cys Tyr Leu Arg Glu Glu Ser Glu Ala Asp Gln Trp Leu
225 230 235 240

Arg Gln Val Trp Gly Glu Val Pro Glu Pro Ser Asp Arg Ser Glu Glu
245 250 255

Pro Glu Thr Pro Ala Ala Tyr Arg Ala Arg Thr
260 265

<210> 143
<211> 164
<212> PRT
<213> Homo sapien

<400> 143

Ala Glu Ala Trp Tyr Gly Ala Arg Phe Pro Val Ser Gly Asp Gly Ser
1 5 10 15

Asp Ala Ile Thr Ala Arg Gln Gly Val Ser Arg Gly Val Lys Ala Ala
20 25 30

Met Asn Arg Val Leu Cys Ala Pro Ala Ala Gly Ala Val Arg Ala Leu
35 40 45

Arg Leu Ile Gly Trp Ala Ser Arg Ser Leu His Pro Leu Pro Gly Ser
50 55 60

Arg Asp Arg Ala His Pro Ala Ala Glu Glu Asp Asp Pro Asp Arg
65 70 75 80

Pro Ile Glu Phe Ser Ser Ser Lys Ala Asn Pro His Arg Trp Ser Val
85 90 95

Gly His Thr Met Gly Lys Gly His Gln Arg Pro Trp Trp Lys Val Leu
100 105 110

Pro Leu Ser Cys Phe Leu Val Ala Leu Ile Ile Trp Cys Tyr Leu Arg
115 120 125

190

Glu Glu Ser Glu Ala Asp Gln Trp Leu Arg Gln Val Trp Gly Glu Val
 130 135 140

Pro Glu Pro Ser Asp Arg Ser Glu Glu Pro Glu Thr Pro Ala Ala Tyr
 145 150 155 160

Arg Ala Arg Thr

<210> 144

<211> 99

<212> PRT

<213> Homo sapien

<400> 144

Met Val Arg Ala Gly Ala Val Gly Ala His Leu Pro Ala Ser Gly Leu
 1 5 10 15

Asp Ile Phe Gly Asp Leu Lys Lys Met Asn Lys Arg Gln Leu Tyr Tyr
 20 25 30

Gln Val Leu Asn Phe Ala Met Ile Val Ser Ser Ala Leu Met Ile Trp
 35 40 45

Lys Gly Leu Ile Val Leu Thr Gly Ser Glu Ser Pro Ile Val Val Val
 50 55 60

Leu Ser Gly Ser Met Glu Pro Ala Phe His Arg Gly Asp Leu Leu Phe
 65 70 80

Leu Thr Asn Phe Arg Glu Asp Pro Ile Arg Ala Glu Ile Met Glu Thr
 85 90 95

Ser Asn Phe

<210> 145

<211> 136

<212> PRT

<213> Homo sapien

<400> 145

Val Ile Cys Glu Arg Glu Leu Gly Val Leu Leu Ala Pro Asp Gln Ser
 1 5 10 15

Arg Glu Ile Gln Leu Leu Ser Ser Pro Phe Pro Glu Leu Pro Pro
 20 25 30

Glu Val Cys Gly Val Thr Arg Cys Ser Met Phe Pro Pro Lys Gly Arg
 35 40 45

Thr Arg Leu Arg Ser Pro Val Ala Ala Leu Pro Arg Ser Pro Gly Ser
 50 55 60

Ser Leu Ala Glu Val Pro Thr Pro Gln His Ser Gly Ser Gly Ser Phe
 65 70 75 80

Leu Pro Ser Gly Ser Phe Leu Ala Gly Gln Cys Pro Arg Leu Ala Arg
 85 90 95

Leu Arg Phe Pro Asp Ala Gln Ala Ser Arg Arg Ser Arg Gly Arg Lys
 100 105 110

Asp Ala Gly Pro Val Gly Gly Arg Gln Val Leu Arg Ser Arg Leu
 115 120 125

Cys His Pro Glu Pro Ala Gly Arg
 130 135

<210> 146
<211> 139
<212> PRT
<213> Homo sapien

<400> 146

Met Ser Lys Thr Phe Arg Gln Thr Glu Gly Ser Gln Gly Asp Arg Arg
 1 5 10 15

Val His Ser Lys Ala Thr Ala Ser Pro Asp Pro Ala Leu Pro Ser Leu
 20 25 30

Leu Trp Thr Gln Glu Lys Ser Asn Pro His Ser Glu Phe Ser His Gln
 35 40 45

Asn Leu Ile Ile Asn Thr Leu Ser Leu Phe Phe Ala Gly Thr Glu Thr
 50 55 60

Thr Ser Thr Thr Leu Arg Tyr Gly Phe Leu Leu Met Leu Lys Tyr Pro
 65 70 75 80

His Val Ala Glu Arg Val Tyr Lys Glu Glu Gln Val Val Gly Pro
 85 90 95

192

His Arg Pro Pro Ala Leu Asp Asp Arg Ala Lys Met Pro Tyr Thr Glu
100 105 110

Ala Val Ile Arg Glu Ile Gln Arg Phe Ala Asp Leu Leu Pro Met Gly
115 120 125

Val Pro His Ile Val Thr Gln His Thr Ser Phe
130 135

<210> 147
<211> 165
<212> PRT
<213> Homo sapien

<400> 147

Arg His Arg Ser Asp Thr Pro Gly Val Trp Cys Gly Gln Asn Thr Pro
1 5 10 15

Asn Ile Pro Asp Leu Leu Pro Ala Pro Leu Lys Gly Leu Arg Glu Gly
20 25 30

Gly Gln Arg Ile Pro Gly Ser Phe Ser Val Pro Thr Ser Val Asp Asn
35 40 45

Gly Ser Asp Ser Leu Gln Leu Pro Ala Ser Glu Arg Pro Ala Ala Ser
50 55 60

Gln Leu Pro Ser Leu Pro Trp His Gln Leu Ser Glu Val Ala Val Gln
65 70 75 80

Met Ser Gly Gly Val Arg Leu Leu Lys Ile Ile Ile Tyr Lys Ile Ile
85 90 95

Tyr Ile Tyr Phe Glu Thr Glu Ser His Ser Val Ala Gln Ala Gly Val
100 105 110

Gln Trp Arg Asp Leu Gly Ser Leu Gln Pro Pro Pro Gly Phe Lys
115 120 125

Lys Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp Tyr Arg Cys Val
130 135 140

Leu Pro Cys Leu Ala Asn Phe Cys Ile Phe Ser Arg Asp Gly Val Ser
145 150 155 160

Pro Cys Trp Pro Gly
165

<210> 148
<211> 136
<212> PRT
<213> Homo sapien

<400> 148

Met Leu Leu Glu Arg Arg Ser Val Met Asp Pro Pro Gly Gln Val Gln
1 5 10 15

Thr Tyr Glu Glu Gly Leu Phe Tyr Ala Gln Lys Ser Lys Lys Pro Leu
20 25 30

Met Val Ile His His Leu Glu Asp Cys Gln Tyr Ser Gln Ala Leu Lys
35 40 45

Lys Val Phe Ala Gln Asn Glu Glu Ile Gln Glu Met Ala Gln Asn Lys
50 55 60

Phe Ile Met Leu Asn Leu Met His Glu Thr Thr Asp Lys Asn Leu Ser
65 70 75 80

Pro Asp Gly Gln Tyr Val Pro Arg Ile Met Phe Val Asp Pro Ser Leu
85 90 95

Thr Val Arg Ala Asp Ile Ala Gly Arg Tyr Ser Asn Arg Leu Tyr Thr
100 105 110

Tyr Glu Pro Arg Asp Leu Pro Leu Ile Glu Asn Met Lys Lys Ala
115 120 125

Leu Arg Leu Ile Gln Ser Glu Leu
130 135

<210> 149
<211> 196
<212> PRT
<213> Homo sapien

<400> 149

Met Glu Gly Asn Gly Pro Ala Ala Val His Tyr Gln Pro Ala Ser Pro
1 5 10 15

Pro Arg Asp Ala Cys Val Tyr Ser Ser Cys Tyr Cys Glu Glu Asn Ile
20 25 30

Trp Lys Leu Cys Glu Tyr Ile Lys Asn His Asp Gln Tyr Pro Leu Glu

194

35

40

45

Glu Cys Tyr Ala Val Phe Ile Ser Asn Glu Arg Lys Met Ile Pro Ile
50 55 60

Trp Lys Gln Gln Ala Arg Pro Gly Asp Gly Pro Val Ile Trp Asp Tyr
65 70 75 80

His Val Val Leu Leu His Val Ser Ser Gly Gly Gln Asn Phe Ile Tyr
85 90 95

Asp Leu Asp Thr Val Leu Pro Phe Pro Cys Leu Phe Asp Thr Tyr Val
100 105 110

Glu Asp Ala Phe Lys Ser Asp Asp Asp Ile His Pro Gln Phe Arg Arg
115 120 125

Lys Phe Arg Val Ile Arg Ala Asp Ser Tyr Leu Lys Asn Phe Ala Ser
130 135 140

Asp Arg Ser His Met Lys Asp Ser Ser Gly Asn Trp Arg Glu Pro Pro
145 150 155 160

Pro Pro Tyr Pro Cys Ile Glu Thr Gly Gly Ile Asn Pro Val Asp Asn
165 170 175

Phe Leu Thr Phe Lys Lys Ile Lys Gly Pro Ser Pro Tyr Tyr Tyr Cys
180 185 190

Leu Ala Phe Ile
195

<210> 150

<211> 69

<212> PRT

<213> Homo sapien

<400> 150

Arg Glu Arg Glu Arg Glu Arg Glu Ser Gly His Lys Asn Cys
1 5 10 15

Phe Val Lys Val Lys Asp Ser Lys Leu Pro Ala Tyr Lys Asp Leu Gly
20 25 30

Lys Asn Leu Pro Phe Pro Thr Tyr Phe Pro Asp Gly Asp Glu Glu Glu
35 40 45

195

Leu Pro Glu Asp Leu Tyr Asp Glu Asn Val Cys Gln Pro Gly Ala Pro
50 55 60

Ser Ile Thr Phe Ala
65

<210> 151
<211> 69
<212> PRT
<213> Homo sapien

<400> 151

Arg Glu Arg Glu Arg Glu Arg Glu Ser Gly His Lys Asn Cys
1 5 10 15

Leu Val Lys Val Lys Asp Ser Lys Leu Pro Ala Tyr Lys Asp Leu Gly
20 25 30

Lys Asn Leu Pro Phe Pro Thr Tyr Phe Pro Asp Gly Asp Glu Glu Glu
35 40 45

Leu Pro Glu Asp Leu Tyr Asp Glu Asn Val Cys Gln Pro Gly Ala Pro
50 55 60

Ser Ile Thr Phe Ala
65

<210> 152
<211> 174
<212> PRT
<213> Homo sapien

<400> 152

Met Glu Ser Arg Thr Leu Leu Gly Gln Leu Trp Val Pro Leu Ala Ser
1 5 10 15

Gly Trp Ala Arg Gly Gln Arg Thr Cys Arg Arg Arg Leu Arg Tyr Gly
20 25 30

Leu Val Lys Val Glu Met Asp Gly Arg Met Asp Ser Leu Gly His Met
35 40 45

Ala Arg Ser Trp Glu Asp Gly His Arg Pro Lys Ser Val Leu Val Tyr
50 55 60

His Cys Thr Ser Gly Asn Leu Asn Pro Cys Asn Arg Gly Lys Met Gly
65 70 75 80

196

Phe Gln Val Leu Ala Thr Phe Glu Ile Pro Ile Pro Phe Glu Arg Ala
85 95

Leu Thr Arg Pro Tyr Ala Asp Phe Thr Thr Ser Asn Phe Arg Thr Gln
100 105 110

Tyr Trp Asn Ala Ile Ser Gln Gln Ala Pro Ala Ile Ile Tyr Asp Phe
115 120 125

Tyr Leu Trp Leu Thr Gly Arg Lys Pro Arg Gln Gly Gln Asp Gly Ser
130 135 140

Lys Ser Asn Gln Pro Pro Leu Gln Pro Ala Thr Ser Cys Trp Gln Asp
145 150 155 160

Leu Phe Leu His Pro Val Lys Ser Gln Gly Gly Thr Arg Ala
165 170

<210> 153
<211> 167
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (44)..(44)
<223> X=any amino acid

<400> 153

Gly Gln Leu Trp Val Pro Leu Ala Ser Gly Trp Ala Arg Gly Gln Arg
1 5 10 15

Thr Cys Arg Arg Leu Arg Tyr Gly Leu Val Lys Val Glu Met Asp
20 25 30

Gly Arg Met Asp Ser Leu Gly His Met Ala Arg Xaa Trp Glu Asp Gly
35 40 45

His Arg Pro Lys Ser Val Leu Val Tyr His Cys Thr Ser Gly Asn Leu
50 55 60

Asn Pro Cys Asn Arg Gly Lys Met Gly Phe Gln Val Leu Ala Thr Phe
65 70 75 80

Glu Ile Pro Ile Pro Phe Glu Arg Ala Leu Thr Arg Pro Tyr Ala Asp
85 90 95

Phe Thr Thr Ser Asn Phe Arg Thr Gln Tyr Trp Asn Ala Ile Ser Gln
100 105 110

Gln Ala Pro Ala Ile Ile Tyr Asp Phe Tyr Leu Trp Leu Thr Gly Arg
115 120 125

Lys Pro Arg Gln Gly Gln Asp Gly Ser Lys Ser Asn Gln Pro Pro Leu
130 135 140

Gln Pro Ala Thr Ser Cys Trp Gln Asp Leu Phe Leu His Pro Val Lys
145 150 155 160

Ser Gln Gly Gly Thr Arg Ala
165

<210> 154
<211> 125
<212> PRT
<213> Homo sapien

<400> 154

Met Gln Gln Ala Arg Glu Thr Ala Val Gln Gln Tyr Lys Lys Leu Glu
1 5 10 15

Glu Glu Ile Gln Thr Leu Arg Val Tyr Tyr Ser Leu His Lys Ser Leu
20 25 30

Ser Gln Glu Glu Asn Leu Lys Asp Gln Phe Asn Tyr Thr Leu Ser Thr
35 40 45

Tyr Glu Glu Ala Leu Lys Asn Arg Glu Asn Ile Val Ser Ile Thr Gln
50 55 60

Gln Gln Asn Glu Glu Leu Ala Thr Gln Leu Gln Gln Ala Leu Thr Glu
65 70 75 80

Arg Ala Asn Met Glu Leu Gln Leu Gln His Ala Arg Glu Ala Ser Gln
85 90 95

Val Ala Asn Glu Lys Val Gln Lys Leu Glu Arg Leu Val Asp Val Leu
100 105 110

Arg Lys Lys Val Gly Thr Gly Thr Met Arg Thr Val Ile
115 120 125

198

<210> 155
<211> 106
<212> PRT
<213> Homo sapien

<400> 155

Met Pro Gln Ser Arg Arg Gln Trp Asp Phe Glu Gly Gly Lys Gly Arg
1 5 10 15

Arg Gln Ala Gly His Ala Leu Arg Gly Ala Arg Thr His Leu Leu His
20 25 30

Pro His Val Phe Arg Ala Leu Ser Leu Trp Glu Ala Phe Phe Arg Thr
35 40 45

Ala Leu Val Asn Trp Lys Arg Asn Pro Ser Pro Trp Trp Pro Cys Ser
50 55 60

Asp Leu Asp Leu Ser Glu Val Thr Leu Pro Leu Arg Ala Leu Gln Ser
65 70 75 80

Leu Leu Ala Gly Gly Thr Ser Pro Ser His Ser His Phe Leu Thr
85 90 95

Leu Ser Leu Cys Ile Thr Gly Ser Leu Leu
100 105

<210> 156
<211> 237
<212> PRT
<213> Homo sapien

<400> 156

Met Pro Gly Pro Ala Pro Gly Arg Gly Ser Gly Val Gly Leu Arg
1 5 10 15

Gly Leu Ser Ser Leu Gln Ala Pro Gln Pro Ser Arg Val Pro Trp Pro
20 25 30

Met Ala Ala Tyr Ser Tyr Arg Pro Gly Pro Gly Ala Gly Pro Gly Pro
35 40 45

Ala Ala Gly Ala Ala Leu Pro Asp Gln Ser Phe Leu Trp Asn Val Phe
50 55 60

Gln Arg Val Asp Lys Asp Arg Ser Gly Val Ile Ser Asp Thr Glu Leu
65 70 75 80

199

Gln Gln Ala Leu Ser Asn Gly Thr Trp Thr Pro Phe Asn Pro Val Thr
85 90 95

Val Arg Ser Ile Ile Ser Met Phe Asp Arg Glu Asn Lys Ala Gly Val
100 105 110

Asn Phe Ser Glu Phe Thr Gly Val Trp Lys Tyr Ile Thr Asp Trp Gln
115 120 125

Asn Val Phe Arg Thr Tyr Asp Arg Asp Asn Ser Gly Met Ile Asp Lys
130 135 140

Asn Glu Leu Lys Gln Ala Leu Ser Gly Phe Gly Tyr Arg Leu Ser Asp
145 150 155 160

Gln Phe His Asp Ile Leu Ile Arg Lys Phe Asp Arg Gln Gly Arg Gly
165 170 175

Gln Ile Ala Phe Asp Asp Phe Ile Gln Gly Cys Ile Val Leu Gln Thr
180 185 190

Leu Ala Pro Ser Pro Arg Pro Glu Cys Gly Ala Asn Thr Ala His
195 200 205

Cys Ser Leu Asp Pro Gln Ala Gln Ala Ile Leu Thr Pro Arg Thr Pro
210 215 220

Lys Val Leu Gly Ser Gln Ala Arg Val Thr Met Leu Ala
225 230 235

<210> 157
<211> 67
<212> PRT
<213> Homo sapien

<400> 157

Lys Asp Gln Ser Ala Ala Glu Asp Pro Ala Arg Ala Arg Thr Arg Ala
1 5 10 15

Arg Arg Arg Ser Ala Lys Glu His Asn Thr His Arg Ala Cys Lys Ala
20 25 30

Ala Ala Arg Ala Pro His Ala Tyr Pro Ala His Thr Val Gln Glu Asp
35 40 45

Asp Val Ala Val His Thr Pro Trp His Gln Pro Thr Pro Arg Thr Ser

200

50

55

60

Ala Ser Leu
65

<210> 158
<211> 156
<212> PRT
<213> Homo sapien

<400> 158

Lys Asp Gln Ser Ala Ala Glu Asp Pro Ala Arg Ala Arg Thr Arg Ala
1 5 10 15

Arg Arg Arg Ser Ala Lys Glu His Asn Thr His Arg Ala Cys Lys Ala
20 25 30

Ala Ala Arg Ala Pro His Ala Tyr Pro Ala His Thr Val Gln Arg Gly
35 40 45

Arg Arg Gly Arg Pro His Pro Val Ala Pro Ala Asn Ala Pro His Leu
50 55 60

Gly Leu Ser Leu Ile Ser Leu Cys Val Val Val Thr Leu Phe Val Ile
65 70 75 80

Val Cys Ser Val Ile Val Cys Tyr Phe Tyr Leu Leu Phe Cys Phe Val
85 90 95

Val Val Cys Val Phe Val Phe Leu Phe Phe Val Phe Leu Phe Phe
100 105 110

Phe Phe Phe Asn Phe Cys Ile Leu Ile Asn Val Phe Asn Tyr Asn Cys
115 120 125

Phe Lys Arg Ile Pro Ala Phe Gln Lys Phe Ile Leu Ser Leu Glu Thr
130 135 140

Arg Gln Gly His Thr Gly Phe Thr Ser Tyr Val Ile
145 150 155

<210> 159
<211> 829
<212> PRT
<213> Homo sapien

<400> 159

201

Met Thr Thr Arg Gln Ala Thr Lys Asp Pro Leu Leu Arg Gly Val Ser
1 5 10 15

Pro Thr Pro Ser Lys Ile Pro Val Arg Ser Gln Lys Arg Thr Pro Phe
20 25 30

Pro Thr Val Thr Ser Cys Ala Val Asp Gln Glu Asn Gln Asp Pro Arg
35 40 45

Arg Trp Val Gln Lys Pro Pro Leu Asn Ile Gln Arg Pro Leu Val Asp
50 55 60

Ser Ala Gly Pro Arg Pro Lys Ala Arg His Gln Ala Glu Thr Ser Gln
65 70 75 80

Arg Leu Val Gly Ile Ser Gln Pro Arg Asn Pro Leu Glu Glu Leu Arg
85 90 95

Pro Ser Pro Arg Gly Gln Asn Val Gly Pro Gly Pro Pro Ala Gln Thr
100 105 110

Glu Ala Pro Gly Thr Ile Glu Phe Val Ala Asp Pro Ala Ala Leu Ala
115 120 125

Thr Ile Leu Ser Gly Glu Gly Val Lys Ser Cys His Leu Gly Arg Gln
130 135 140

Pro Ser Leu Ala Lys Arg Val Leu Val Arg Gly Ser Gln Gly Thr
145 150 155 160

Thr Gln Arg Val Gln Gly Val Arg Ala Ser Ala Tyr Leu Ala Pro Arg
165 170 175

Thr Pro Thr His Arg Leu Asp Pro Ala Arg Ala Ser Cys Phe Ser Arg
180 185 190

Leu Glu Gly Pro Gly Pro Arg Gly Arg Thr Leu Cys Pro Gln Arg Leu
195 200 205

Gln Ala Leu Ile Ser Pro Ser Gly Pro Ser Phe His Pro Ser Thr Arg
210 215 220

Pro Ser Phe Gln Glu Leu Arg Arg Glu Thr Ala Gly Ser Ser Arg Thr
225 230 235 240

Ser Val Ser Gln Ala Ser Gly Leu Leu Leu Glu Thr Pro Val Gln Pro

202
245 250 255

Ala Phe Ser Leu Pro Lys Gly Glu Arg Glu Val Val Thr His Ser Asp
260 265 270

Glu Gly Gly Val Ala Ser Leu Gly Leu Ala Gln Arg Val Pro Leu Arg
275 280 285

Glu Asn Arg Glu Met Ser His Thr Arg Asp Ser His Asp Ser His Leu
290 295 300

Met Pro Ser Pro Ala Pro Val Ala Gln Pro Leu Pro Gly His Val Val
305 310 315 320

Pro Cys Pro Ser Pro Phe Gly Arg Ala Gln Arg Val Pro Ser Pro Gly
325 330 335

Pro Pro Thr Leu Thr Ser Tyr Ser Val Leu Arg Arg Leu Thr Val Gln
340 345 350

Pro Lys Thr Arg Phe Thr Pro Met Pro Ser Thr Pro Arg Val Gln Gln
355 360 365

Ala Gln Trp Leu Arg Gly Val Ser Pro Gln Ser Cys Ser Glu Asp Pro
370 375 380

Ala Leu Pro Trp Glu Gln Val Ala Val Arg Leu Phe Asp Gln Glu Ser
385 390 395 400

Cys Ile Arg Ser Leu Glu Gly Ser Gly Lys Pro Pro Val Ala Thr Pro
405 410 415

Ser Gly Pro His Ser Asn Arg Thr Pro Ser Leu Gln Glu Val Lys Ile
420 425 430

Gln Val Ser Leu Cys Gly Gln Gln Leu Cys Cys Leu Leu Asn Ser Asp
435 440 445

Trp Ala Glu Glu Glu Gly Lys Glu Met Gly Asp Gln Glu Glu Asp Ser
450 455 460

Val Gly Arg Leu Leu Asn Ala His Leu Asp Val Thr Leu Gly Cys Ser
465 470 475 480

Leu Pro Pro Gln Arg Ile Gly Ile Leu Gln Gln Leu Leu Arg Gln Glu
485 490 495

Val Glu Gly Leu Val Gly Gly Gln Cys Val Pro Leu Asn Gly Gly Ser
500 505 510

Ser Leu Asp Met Val Glu Leu Gln Pro Leu Leu Thr Glu Ile Ser Arg
515 520 525

Thr Leu Asn Ala Thr Glu His Asn Ser Gly Thr Ser His Leu Pro Gly
530 535 540

Leu Leu Lys His Ser Gly Leu Pro Lys Pro Cys Leu Pro Glu Glu Cys
545 550 555 560

Gly Glu Pro Gln Pro Cys Pro Pro Ala Glu Pro Gly Pro Pro Glu Ala
565 570 575

Phe Cys Arg Ser Glu Pro Glu Ile Pro Glu Pro Ser Leu Gln Glu Gln
580 585 590

Leu Glu Val Pro Glu Pro Tyr Pro Pro Ala Glu Pro Arg Pro Leu Glu
595 600 605

Ser Cys Cys Arg Ser Glu Pro Glu Ile Pro Glu Ser Ser Arg Gln Glu
610 615 620

Gln Leu Glu Val Pro Glu Pro Cys Pro Pro Ala Glu Pro Arg Pro Leu
625 630 635 640

Glu Ser Tyr Cys Arg Ile Glu Pro Glu Ile Pro Glu Ser Ser Arg Gln
645 650 655

Glu Gln Leu Glu Val Pro Glu Pro Cys Pro Pro Ala Glu Pro Gly Pro
660 665 670

Leu Gln Pro Ser Thr Gln Gly Gln Ser Gly Pro Pro Gly Pro Cys Pro
675 680 685

Arg Val Glu Leu Gly Ala Ser Glu Pro Cys Thr Leu Glu His Arg Ser
690 695 700

Leu Glu Ser Ser Leu Pro Pro Cys Cys Ser Gln Trp Ala Pro Ala Thr
705 710 715 720

Thr Ser Leu Ile Phe Ser Ser Gln His Pro Leu Cys Ala Ser Pro Pro
725 730 735

204

Ile Cys Ser Leu Gln Ser Leu Arg Pro Pro Ala Gly Gln Ala Gly Leu
740 745 750

Ser Asn Leu Ala Pro Arg Thr Leu Ala Leu Arg Glu Arg Leu Lys Ser
755 760 765

Cys Leu Thr Ala Ile His Cys Phe His Glu Ala Arg Leu Asp Asp Glu
770 775 780

Cys Ala Phe Tyr Thr Ser Arg Ala Pro Pro Ser Gly Pro Thr Arg Val
785 790 795 800

Cys Thr Asn Pro Val Ala Thr Leu Leu Glu Trp Gln Asp Ala Leu Cys
805 810 815

Phe Ile Pro Val Gly Ser Ala Ala Pro Gln Gly Ser Pro
820 825

<210> 160

<211> 443

<212> PRT

<213> Homo sapien

<400> 160

Ala Ile Met Thr Thr Arg Gln Ala Thr Lys Asp Pro Leu Leu Arg Gly
1 5 10 15

Val Ser Pro Thr Pro Ser Lys Ile Pro Val Arg Ser Gln Lys Arg Thr
20 25 30

Pro Phe Pro Thr Val Thr Ser Cys Ala Val Asp Gln Glu Asn Gln Asp
35 40 45

Pro Arg Arg Trp Val Gln Lys Pro Pro Leu Asn Ile Gln Arg Pro Leu
50 55 60

Val Asp Ser Ala Gly Pro Arg Pro Lys Ala Arg His Gln Ala Glu Thr
65 70 75 80

Ser Gln Arg Leu Val Gly Ile Ser Gln Pro Arg Asn Pro Leu Glu Glu
85 90 95

Leu Arg Pro Ser Pro Arg Gly Gln Asn Val Gly Pro Gly Pro Pro Ala
100 105 110

Gln Thr Glu Ala Pro Gly Thr Ile Glu Phe Val Ala Asp Pro Ala Ala

205

115

120

125

Leu Ala Thr Ile Leu Ser Gly Glu Gly Val Lys Ser Cys His Leu Gly
130 135 140

Arg Gln Pro Ser Leu Ala Lys Arg Val Leu Val Arg Gly Ser Gln Gly
145 150 155 160

Gly Thr Thr Gln Arg Val Gln Gly Val Arg Ala Ser Ala Tyr Leu Ala
165 170 175

Pro Arg Thr Pro Thr His Arg Leu Asp Pro Ala Arg Ala Ser Cys Phe
180 185 190

Ser Arg Leu Glu Gly Pro Gly Pro Arg Gly Arg Thr Leu Cys Pro Gln
195 200 205

Arg Leu Gln Ala Leu Ile Ser Pro Ser Gly Pro Ser Phe His Pro Ser
210 215 220

Thr Arg Pro Ser Phe Gln Glu Leu Arg Arg Glu Thr Ala Gly Ser Ser
225 230 235 240

Arg Thr Ser Val Ser Gln Ala Ser Gly Leu Leu Leu Glu Thr Pro Val
245 250 255

Gln Pro Ala Phe Ser Leu Pro Lys Gly Glu Arg Glu Val Val Thr His
260 265 270

Ser Asp Glu Gly Gly Val Ala Ser Leu Gly Leu Ala Gln Arg Val Pro
275 280 285

Leu Arg Glu Asn Arg Glu Met Ser His Thr Arg Asp Ser His Asp Ser
290 295 300

His Leu Met Pro Ser Pro Ala Pro Val Ala Gln Pro Leu Pro Gly His
305 310 315 320

Val Val Pro Cys Pro Ser Pro Phe Gly Arg Ala Gln Arg Val Pro Ser
325 330 335

Pro Gly Pro Pro Thr Leu Thr Ser Tyr Ser Val Leu Arg Arg Leu Thr
340 345 350

Val Gln Pro Lys Thr Arg Phe Thr Pro Met Pro Ser Thr Pro Arg Val
355 360 365

Gln Gln Ala Gln Trp Leu Arg Gly Val Ser Pro Gln Ser Cys Ser Glu
370 375 380

Asp Pro Ala Leu Pro Trp Glu Gln Val Ala Val Arg Leu Phe Asp Gln
385 390 395 400

Glu Ser Cys Ile Arg Ser Leu Glu Gly Ser Gly Lys Pro Pro Val Ala
405 410 415

Thr Pro Ser Gly Pro His Ser Asn Arg Thr Pro Ser Leu Gln Glu Val
420 425 430

Lys Ile Gln Val Ser Leu Cys Gly Gln Gln Leu
435 440

<210> 161
<211> 138

<212> PRT
<213> Homo sapien

<400> 161

Met Leu Pro His Leu Pro Pro Trp Pro Ser Leu Ala Leu Pro Gln Glu
1 5 10 15

Glu Gly Arg Gly Cys Thr Ser Ser Pro Val Leu Leu Ile Gly Leu Ala
20 25 30

Val Gly Gly Gly Gly Glu Asp Ser Thr Trp Trp Lys Tyr Arg Thr
35 40 45

Pro Asp Leu Pro Leu Asn Phe Pro Cys Pro Ser Gly Leu Ser Asn Leu
50 55 60

Ala Pro Arg Thr Leu Ala Leu Arg Glu Arg Leu Lys Ser Cys Leu Thr
65 70 75 80

Ala Ile His Cys Phe His Glu Ala Arg Leu Asp Asp Glu Cys Ala Phe
85 90 95

Tyr Thr Ser Arg Ala Pro Pro Ser Gly Pro Thr Arg Val Cys Thr Asn
100 105 110

Pro Val Ala Thr Leu Leu Glu Trp Gln Asp Ala Leu Cys Phe Ile Pro
115 120 125

207

Val Gly Ser Ala Ala Pro Gln Gly Ser Pro
130 135

<210> 162
<211> 60
<212> PRT
<213> Homo sapien

<400> 162

Met Arg Ala Arg Thr Pro Pro Ala Ala Pro Lys Glu Lys Ala Phe Ser
1 5 10 15

Ser Glu Ile Glu Asp Leu Pro Tyr Leu Ser Thr Thr Glu Met Tyr Leu
20 25 30

Cys Arg Trp His Gln Pro Pro Ser Pro Leu Pro Leu Arg Glu Ser
35 40 45

Ser Pro Lys Lys Glu Glu Thr Val Ala Ser Lys Ala
50 55 60

<210> 163
<211> 99
<212> PRT
<213> Homo sapien

<400> 163

Lys Lys Gly Phe Leu Cys Cys Glu Met His Arg Thr Ile Leu Cys His
1 5 10 15

Ala Arg Leu Phe Leu Gln Leu Ile Leu Cys Glu Ile Trp Glu Gly Gly
20 25 30

Leu Trp Val Phe Ser Gly Ala Asn Gly Asn Phe Trp Val Gly Glu Pro
35 40 45

Ala Trp Gly Gly Glu Phe Ser Pro Gly Pro Pro Leu Phe Asn Tyr Ile
50 55 60

Asn Ile Tyr Leu Tyr Ile Tyr Val Pro Val Trp Gly Ala Gly Gly Ile
65 70 75 80

Cys Gln Arg Pro Thr Val Leu Leu Tyr Leu Thr Ile Leu His Lys Gly
85 90 95

Ser Lys Met

<210> 164
<211> 294
<212> PRT
<213> Homo sapien

<400> 164

Met Phe Phe Ser Ala Ala Leu Arg Ala Arg Ala Ala Gly Leu Thr Ala
1 5 10 15

His Trp Gly Arg His Val Arg Asn Leu His Lys Thr Ala Met Gln Asn
20 25 30

Gly Ala Gly Gly Ala Leu Phe Val His Arg Asp Thr Pro Glu Asn Asn
35 40 45

Pro Asp Thr Pro Phe Asp Phe Thr Pro Glu Asn Tyr Lys Arg Ile Glu
50 55 60

Ala Ile Val Lys Asn Tyr Pro Glu Gly His Lys Ala Ala Ala Val Leu
65 70 75 80

Pro Val Leu Asp Leu Ala Gln Arg Gln Asn Gly Trp Leu Pro Ile Ser
85 90 95

Ala Met Asn Lys Val Ala Glu Val Leu Gln Val Pro Pro Met Arg Val
100 105 110

Tyr Glu Val Ala Thr Phe Tyr Thr Met Tyr Asn Arg Lys Pro Val Gly
115 120 125

Lys Tyr His Ile Gln Val Cys Thr Thr Pro Cys Met Leu Arg Asn
130 135 140

Ser Asp Ser Ile Leu Glu Ala Ile Gln Lys Lys Leu Gly Ile Lys Val
145 150 155 160

Gly Glu Thr Thr Pro Asp Lys Leu Phe Thr Leu Ile Glu Val Glu Cys
165 170 175

Leu Gly Ala Cys Val Asn Ala Pro Met Val Gln Ile Asn Asp Asn Tyr
180 185 190

Tyr Glu Asp Leu Thr Ala Lys Asp Ile Glu Glu Ile Ile Asp Glu Leu
195 200 205

Lys Ala Gly Lys Ile Pro Lys Pro Gly Pro Arg Ser Gly Arg Phe Ser

209

210

215

220

Cys Glu Pro Ala Gly Gly Leu Thr Ser Leu Thr Glu Pro Pro Lys Gly
 225 230 235 240

Pro Gly Phe Gly Val Gln Cys Val His Leu His Arg Lys Phe Gln Gly
 245 250 255

Ala Ile Ala Val Val Val Asn His Arg Ile Ser Val Gly Met Ala Glu
 260 265 270

Gly Glu Thr Gly Leu Gly Cys Arg Glu Leu Val Glu Val Val Gln Pro
 275 280 285

Tyr Leu Pro Gly Arg Pro
 290

<210> 165
<211> 250
<212> PRT
<213> Homo sapien

<400> 165

Met Phe Phe Ser Ala Ala Leu Arg Ala Arg Ala Ala Gly Leu Thr Ala
 1 5 10 15

His Trp Gly Arg His Val Arg Asn Leu His Lys Thr Ala Met Gln Asn
 20 25 30

Gly Ala Gly Gly Ala Leu Phe Val His Arg Asp Thr Pro Glu Asn Asn
 35 40 45

Pro Asp Thr Pro Phe Asp Phe Thr Pro Glu Asn Tyr Lys Arg Ile Glu
 50 55 60

Ala Ile Val Lys Asn Tyr Pro Glu Gly His Lys Ala Ala Ala Val Leu
 65 70 75 80

Pro Val Leu Asp Leu Ala Gln Arg Gln Asn Gly Trp Leu Pro Ile Ser
 85 90 95

Ala Met Asn Lys Val Ala Glu Val Leu Gln Val Pro Pro Met Arg Val
 100 105 110

Tyr Glu Val Ala Thr Phe Tyr Thr Met Tyr Asn Arg Lys Pro Val Gly
 115 120 125

210

Lys Tyr His Ile Gln Val Cys Thr Thr Thr Pro Cys Met Leu Arg Asn
130 135 140

Ser Asp Ser Ile Leu Glu Ala Ile Gln Lys Lys Leu Gly Ile Lys Val
145 150 155 160

Gly Glu Thr Thr Pro Asp Lys Leu Phe Thr Leu Ile Glu Val Glu Cys
165 170 175

Leu Gly Ala Cys Val Asn Ala Pro Met Val Gln Ile Asn Asp Asn Tyr
180 185 190

Tyr Glu Asp Leu Thr Ala Lys Asp Ile Glu Glu Ile Ile Asp Glu Leu
195 200 205

Lys Ala Gly Lys Ile Pro Lys Pro Gly Pro Arg Ser Gly Arg Phe Ser
210 215 220

Cys Glu Pro Ala Gly Gly Leu Thr Ser Leu Thr Glu Arg Pro Pro Val
225 230 235 240

Cys Cys Gln Ser Phe Glu Ala Cys Arg Val
245 250

<210> 166

<211> 232

<212> PRT

<213> Homo sapien

<400> 166

Met Phe Phe Ser Ala Ala Leu Arg Ala Arg Ala Ala Gly Leu Thr Ala
1 5 10 15

His Trp Gly Arg His Val Arg Asn Leu His Lys Thr Ala Met Gln Asn
20 25 30

Gly Ala Gly Gly Ala Leu Phe Val His Arg Asp Thr Pro Glu Asn Asn
35 40 45

Pro Asp Thr Pro Phe Asp Phe Thr Pro Glu Asn Tyr Lys Arg Ile Glu
50 55 60

Ala Ile Val Lys Asn Tyr Pro Glu Gly His Lys Ala Ala Ala Val Leu
65 70 75 80

Pro Val Leu Asp Leu Ala Gln Arg Gln Asn Gly Trp Leu Pro Ile Ser

211

85

90

95

Ala Met Asn Lys Val Ala Glu Val Leu Gln Val Pro Pro Met Arg Val
100 105 110

Tyr Glu Val Ala Thr Phe Tyr Thr Met Tyr Asn Arg Lys Pro Val Gly
115 120 125

Lys Tyr His Ile Gln Val Cys Thr Thr Pro Cys Met Leu Arg Asn
130 135 140

Ser Asp Ser Ile Leu Glu Ala Ile Gln Lys Lys Leu Gly Arg Glu Tyr
145 150 155 160

Met Ile Phe Val Thr Leu Ile Lys Ser Arg Ile Val Ser Leu Asp Leu
165 170 175

Val His Phe Tyr Leu Lys Phe Pro Thr Ser Ala Ile Leu Leu Asp Leu
180 185 190

Tyr Leu Pro Ser Asn Ile Leu Cys Tyr Cys Val Ser Thr Ser Leu Phe
195 200 205

Leu Pro Ile Trp Tyr Ser Ser Val Leu Ser Val Lys Ala Glu Phe
210 215 220

Leu Ile Phe Ser Phe Leu Ile Ser
225 230

<210> 167
<211> 28
<212> PRT
<213> Homo sapien

<400> 167

Met Asp Ser Arg Pro Arg Tyr Ile Pro Phe Lys Gln Tyr Ala Gly Lys
1 5 10 15

Tyr Val Leu Leu Ser Thr Trp Pro Ala Thr Glu Ala
20 25

<210> 168
<211> 106
<212> PRT
<213> Homo sapien

<400> 168

212

Trp Ile Arg Gly Arg Gly Thr Ser Pro Ser Ser Ser Met Leu Ala Asn
1 5 10 15

Thr Ser Ser Cys Gln Arg Gly Gln Leu Leu Arg Pro Asp Gly Pro Val
20 25 30

His Gln Val Asp Arg Leu Cys Gly Ala Cys Pro Gly Gln Arg Val Phe
35 40 45

Leu Cys Pro Gly Glu Pro Gly Ala Lys Ser Gly Arg His Leu Ser Gly
50 55 60

Gly Val Pro Pro Tyr Thr Glu Cys Asp His Ala Gln Pro Leu Ala Arg
65 70 75 80

Pro Gly Ala Val Glu Ser Cys Asn His Glu Val Cys Ala Gln Thr Gly
85 90 95

Glu Thr Val Gln Pro Leu Met Ala Arg Arg
100 105

<210> 169

<211> 137

<212> PRT

<213> Homo sapien

<400> 169

Met Lys Val Leu Gly Arg Ser Phe Phe Trp Val Leu Phe Pro Val Leu
1 5 10 15

Pro Trp Ala Val Gln Ala Val Glu His Glu Glu Val Ala Gln Arg Val
20 25 30

Ile Lys Leu His Arg Gly Arg Gly Val Ala Ala Met Gln Ser Arg Gln
35 40 45

Trp Val Arg Asp Ser Cys Arg Lys Leu Ser Gly Leu Leu Arg Gln Lys
50 55 60

Asn Ala Val Leu Asn Lys Leu Lys Thr Ala Ile Gly Ala Val Glu Lys
65 70 75 80

Asp Val Gly Leu Ser Asp Glu Glu Lys Leu Phe Gln Val His Thr Phe
85 90 95

Glu Ile Phe Gln Lys Glu Leu Asn Glu Ser Glu Asn Ser Val Phe Gln
100 105 110

Ala Val Tyr Gly Leu Gln Arg Ala Leu Gln Gly Asp Tyr Asn Asp Gly
115 120 125

Pro Trp Lys Gly Ser Val Cys Gly Glu
130 135

<210> 170
<211> 241
<212> PRT
<213> Homo sapien

<400> 170

Met Lys Val Leu Gly Arg Ser Phe Phe Trp Val Leu Phe Pro Val Leu
1 5 10 15

Pro Trp Ala Val Gln Ala Val Glu His Glu Glu Val Ala Gln Arg Val
20 25 30

Ile Lys Leu His Arg Gly Arg Gly Val Ala Ala Met Gln Ser Arg Gln
35 40 45

Trp Val Arg Asp Ser Cys Arg Lys Leu Ser Gly Leu Leu Arg Gln Lys
50 55 60

Asn Ala Val Leu Asn Lys Leu Lys Thr Ala Ile Gly Ala Val Glu Lys
65 70 75 80

Asp Val Gly Leu Ser Asp Glu Glu Lys Leu Phe Gln Val His Thr Phe
85 90 95

Glu Ile Phe Gln Lys Glu Leu Asn Glu Ser Glu Asn Ser Val Phe Gln
100 105 110

Ala Val Tyr Gly Leu Gln Arg Ala Leu Gln Gly Asp Tyr Lys Asp Val
115 120 125

Val Asn Met Lys Glu Ser Ser Arg Gln Arg Leu Glu Ala Leu Arg Glu
130 135 140

Ala Ala Ile Lys Glu Glu Thr Glu Tyr Met Glu Leu Leu Ala Ala Glu
145 150 155 160

Lys His Gln Val Glu Ala Leu Lys Asn Met Gln His Gln Asn Gln Ser
165 170 175

214

Leu Ser Met Leu Asp Glu Ile Leu Glu Asp Val Arg Lys Ala Ala Asp
180 185 190

Arg Leu Glu Glu Glu Ile Glu Glu His Ala Phe Asp Asp Asn Lys Ser
195 200 205

Val Ser Val Pro Glu Gln Leu Leu Leu His Leu Leu Ser His Ser Leu
210 215 220

Ile Arg Arg His Val Val Glu Ile Val His Val Tyr Val Phe Asn Val
225 230 235 240

Asp

<210> 171

<211> 102

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (15)..(15)

<223> X=any amino acid

<400> 171

Trp Val Ile Gly Phe Ser Pro Leu Arg Pro Thr His Cys Thr Xaa Thr
1 5 10 15

Leu Arg Asp Pro Arg Gly Ala Gly Ala Asp Val Arg Ser Ala Pro Ser
20 25 30

Arg Gly Gly Arg Ala Gly Gln Trp Gly Pro His Arg Gly Gly Val Leu
35 40 45

Val Ser Gly Pro Gly Trp Arg Thr Arg Thr Leu Val Pro Arg Ala Gly
50 55 60

Arg Arg Trp Val His Gly Arg Pro His Pro Arg Ile Pro Ser Pro Ala
65 70 75 80

Pro Ser Leu Asp Ser Pro Val Asn Pro Ala Ala Ser Arg Arg Pro Thr
85 90 95

Trp Ser Trp Pro Val Leu
100

215

<210> 172
<211> 207
<212> PRT
<213> Homo sapien

<400> 172

Met Lys Ser Ser Gly His Arg Glu Trp Gly Val Gly Lys Pro Gly Thr
1 5 10 15

Pro Gly Asp Arg Ala Arg Glu Gly Gly Ser Gly Pro Asp Pro Ala Pro
20 25 30

Ala Arg Gly Ala Ser Ser Gly Ala Ala Leu Arg Gly Gln Asn Val Ala
35 40 45

Val Ala Glu Thr Arg Arg Gly Arg Pro Asn Ala Thr Leu Gly Pro Ser
50 55 60

Pro Leu Gln Arg Pro Arg Pro Val Thr Cys Pro Arg Phe Ala Ser His
65 70 75 80

Pro Glu Ala Gly Ala Arg Ala Glu Pro Ala Ala Met Ser Gly Glu Pro
85 90 95

Gly Gln Thr Ser Val Ala Pro Pro Pro Glu Glu Val Glu Pro Gly Ser
100 105 110

Gly Val Arg Ile Val Val Glu Tyr Cys Glu Pro Cys Gly Phe Glu Ala
115 120 125

Thr Tyr Leu Glu Leu Ala Ser Ala Val Lys Glu Gln Tyr Pro Gly Ile
130 135 140

Glu Ile Glu Ser Arg Leu Gly Gly Thr Gly Ala Phe Glu Ile Glu Ile
145 150 155 160

Asn Gly Gln Leu Val Phe Ser Lys Leu Glu Asn Gly Gly Phe Pro Tyr
165 170 175

Glu Lys Asp Val Ser Ile Tyr Ser Val Gly Arg Thr Ser Trp Ser Pro
180 185 190

Tyr Pro Asn Ser Ala Ser Ser Cys His Ser Thr Pro Leu Ala His
195 200 205

<210> 173
<211> 208

216

<212> PRT

<213> Homo sapien

<400> 173

Ser His Glu Val Gln Arg Thr Pro Gly Val Gly Ser Gly Glu Ala Arg
1 5 10 15

His Ser Gly Arg Pro Gly Gln Gly Arg Arg Val Trp Thr Gly Pro Ser
20 25 30

Pro Cys Pro Gly Ser Glu Leu Arg Ser Cys Pro Thr Arg Ser Lys Arg
35 40 45

Ser Ser Gly Gly Asp Pro Gln Gly Ala Pro Glu Arg His Pro Arg Pro
50 55 60

Leu Pro Ala Pro Glu Ala Pro Pro Arg His Val Pro Ala Val Arg Val
65 70 75 80

Thr Pro Gly Ser Arg Gly Pro Ser Gly Pro Ala Ala Met Ser Gly Glu
85 90 95

Pro Gly Gln Thr Ser Val Ala Pro Pro Pro Glu Glu Val Glu Pro Gly
100 105 110

Ser Gly Val Arg Ile Val Val Glu Tyr Cys Glu Pro Cys Gly Phe Glu
115 120 125

Ala Thr Tyr Leu Glu Leu Ala Ser Ala Val Lys Glu Gln Tyr Pro Gly
130 135 140

Ile Glu Ile Glu Ser Arg Leu Gly Gly Thr Gly Ala Phe Glu Ile Glu
145 150 155 160

Ile Asn Gly Gln Leu Val Phe Ser Lys Leu Glu Asn Gly Gly Phe Pro
165 170 175

Tyr Glu Lys Asp Val Ser Ile Tyr Ser Val Gly Arg Thr Ser Trp Ser
180 185 190

Pro Tyr Pro Asn Ser Ala Ser Ser Cys His Ser Thr Pro Leu Ala His
195 200 205

<210> 174

<211> 267

<212> PRT

<213> Homo sapien

<400> 174

Met Val Ser Asn Ser Ala Gly Ser Asn Ser Arg Gln Leu Pro Leu Pro
1 5 10 15

Leu Ser Ala Asp Ala Pro Pro Ala Ser Ser Ser His Trp Ser Trp Gln
20 25 30

Pro Ser Arg His Thr Asn Gln Pro Ile Asp Arg Ala Ile Leu Arg Ser
35 40 45

Arg Pro Cys Cys Arg Leu Ser Arg Thr Cys His Trp Ser Leu Gln Pro
50 55 60

Pro Pro Pro Pro Ala Arg Gln Trp Leu Gly Gly Leu Ala Gly Ala
65 70 75 80

Gly Arg Ser Ser Cys Ala Cys Ala Leu Gly Leu Pro Ser Ala Gly Cys
85 90 95

Ser Ala Gly Arg Ala Arg Leu Arg Gly Ala Ala Leu Glu Glu Thr Glu
100 105 110

Ala Ala Gly Gly Pro Glu Ala Gln Glu Glu Asp Glu Asp Glu Glu Glu
115 120 125

Ala Leu Pro His Ser Glu Ala Met Asp Val Phe Gln Glu Gly Leu Ala
130 135 140

Met Val Val Gln Asp Pro Leu Leu Cys Asp Leu Pro Ile Gln Val Thr
145 150 155 160

Leu Glu Glu Val Asn Ser Gln Ile Ala Leu Glu Tyr Gly Gln Ala Met
165 170 175

Thr Val Arg Val Cys Lys Met Asp Gly Glu Val Met Pro Val Val Val
180 185 190

Val Gln Ser Ala Thr Val Leu Asp Leu Lys Lys Ala Ile Gln Arg Tyr
195 200 205

Val Gln Leu Lys Gln Glu Arg Glu Gly Gly Ile Gln His Ile Ser Trp
210 215 220

Ser Tyr Val Trp Arg Thr Tyr His Leu Thr Ser Ala Gly Glu Lys Leu
225 230 235 240

Thr Glu Asp Arg Lys Lys Leu Arg Asp Tyr Gly Ile Arg Asn Arg Asp
245 250 255

Glu Val Ser Phe Ile Lys Lys Leu Arg Gln Lys
260 265

<210> 175
<211> 225
<212> PRT
<213> Homo sapien

<400> 175

Thr Gly Arg Phe Cys Ala Pro Gly Leu Leu Gln Ala Val Ser His Leu
1 5 10 15

Ser Leu Val Thr Ala Ala Ala Pro Pro Pro Arg Arg Ala Ser Gly Trp
20 25 30

Ala Ala Ser Leu Gly Arg Ala Ala Val Pro Ala Arg Ala Arg Leu Ala
35 40 45

Ser Leu Val Arg Ala Gly Ser Ala Gly Arg Ala Arg Leu Arg Gly Ala
50 55 60

Ala Leu Glu Glu Thr Glu Ala Ala Gly Gly Pro Glu Ala Gln Glu Glu
65 70 75 80

Asp Glu Asp Glu Glu Ala Leu Pro His Ser Glu Ala Met Asp Val
85 90 95

Phe Gln Glu Gly Leu Ala Met Val Val Gln Asp Pro Leu Leu Cys Asp
100 105 110

Leu Pro Ile Gln Val Thr Leu Glu Glu Val Asn Ser Gln Ile Ala Leu
115 120 125

Glu Tyr Gly Gln Ala Met Thr Val Arg Val Cys Lys Met Asp Gly Glu
130 135 140

Val Met Pro Val Val Val Gln Ser Ala Thr Val Leu Asp Leu Lys
145 150 155 160

Lys Ala Ile Gln Arg Tyr Val Gln Leu Lys Gln Glu Arg Glu Gly Gly
165 170 175

219

Ile Gln His Ile Ser Trp Ser Tyr Val Trp Arg Thr Tyr His Leu Thr
180 185 190

Ser Ala Gly Glu Lys Leu Thr Glu Asp Arg Lys Lys Leu Arg Asp Tyr
195 200 205

Gly Ile Arg Asn Arg Asp Glu Val Ser Phe Ile Lys Lys Leu Arg Gln
210 215 220

Lys
225

<210> 176
<211> 224
<212> PRT
<213> Homo sapien

<400> 176

Met Val Ser Asn Ser Ala Gly Ser Asn Ser Arg Gln Leu Pro Leu Pro
1 5 10 15

Leu Ser Ala Asp Ala Pro Pro Ala Ser Ser Ser His Trp Ser Trp Gln
20 25 30

Pro Ser Arg His Thr Asn Gln Pro Ile Asp Arg Ala Ile Leu Arg Ser
35 40 45

Arg Pro Cys Cys Arg Leu Ser Arg Thr Cys His Trp Ser Leu Gln Pro
50 55 60

Pro His Pro Pro Arg Arg Ala Ser Gly Trp Ala Ala Ser Leu Gly Arg
65 70 75 80

Ala Ala Val Pro Ala Arg Ala Arg Leu Ala Ser Leu Val Arg Ala Gly
85 90 95

Ser Ala Gly Arg Ala Arg Leu Arg Gly Ala Ala Leu Glu Glu Thr Glu
100 105 110

Ala Ala Gly Gly Pro Glu Ala Gln Glu Glu Asp Glu Asp Glu Glu Glu
115 120 125

Ala Leu Pro His Ser Glu Ala Met Asp Val Phe Gln Glu Gly Leu Ala
130 135 140

Met Val Val Gln Asp Pro Leu Leu Cys Asp Leu Pro Ile Gln Val Thr
145 150 155 160

220

Leu Glu Glu Val Asn Ser Gln Ile Ala Leu Glu Tyr Gly Gln Ala Met
165 170 175

Thr Val Arg Val Cys Lys Met Asp Gly Glu Val Met Pro Val Val Val
180 185 190

Val Gln Ser Ala Thr Val Leu Asp Leu Lys Lys Ala Ile Gln Arg Tyr
195 200 205

Val Gln Leu Lys Gln Glu Arg Glu Gly Gly Ile Gln His Ile Ser Trp
210 215 220

<210> 177

<211> 300

<212> PRT

<213> Homo sapien

<400> 177

Met Val Ser Asn Ser Ala Gly Ser Asn Ser Arg Gln Leu Pro Leu Pro
1 5 10 15

Leu Ser Ala Asp Ala Pro Pro Ala Ser Ser Ser His Trp Ser Trp Gln
20 25 30

Pro Ser Arg His Thr Asn Gln Pro Ile Asp Arg Ala Ile Leu Arg Ser
35 40 45

Arg Pro Cys Cys Arg Leu Ser Arg Thr Cys His Trp Ser Leu Gln Pro
50 55 60

Pro His Pro Pro Arg Arg Ala Ser Gly Trp Ala Ala Ser Leu Gly Arg
65 70 75 80

Ala Ala Val Pro Ala Arg Ala Arg Leu Ala Ser Leu Val Arg Ala Gly
85 90 95

Ser Ala Gly Arg Ala Arg Leu Arg Gly Ala Ala Leu Glu Glu Thr Glu
100 105 110

Ala Ala Gly Gly Pro Glu Ala Gln Glu Glu Asp Glu Asp Glu Glu Glu
115 120 125

Ala Leu Pro His Ser Glu Ala Met Asp Val Phe Gln Glu Gly Leu Ala
130 135 140

221

Met Val Val Gln Asp Pro Leu Leu Cys Asp Leu Pro Ile Gln Val Thr
145 150 155 160

Leu Glu Glu Val Asn Ser Gln Ile Ala Leu Glu Tyr Gly Gln Ala Met
165 170 175

Thr Val Arg Val Cys Lys Met Asp Gly Glu Val Met Arg Lys Cys Tyr
180 185 190

Pro Pro Pro Phe Arg Phe Met Trp Ser Arg Leu Ser Gln Gln Glu Asp
195 200 205

Leu Thr Val Leu Val Ser Leu Leu Arg Asn Ser Gln Ala Met Pro Arg
210 215 220

Gly Thr Gly Ala Thr Thr Asn Leu Pro Cys Ala Gln Arg Cys Trp Phe
225 230 235 240

Leu Ser Cys His Arg Arg Leu Trp Leu Trp Val Leu Thr Met Asp Leu
245 250 255

Leu Pro Ser Val Ser Val Val Ala Ala Val Val Val Gln Ser Ala
260 265 270

Thr Val Leu Asp Leu Lys Lys Ala Ile Gln Arg Tyr Val Gln Leu Lys
275 280 285

Gln Glu Arg Glu Gly Gly Ile Gln His Ile Ser Trp
290 295 300

<210> 178
<211> 236
<212> PRT
<213> Homo sapien

<400> 178

Gly His Val Leu Gln Ala Lys Arg Trp Gln Arg Cys Pro Ser Ser Thr
1 5 10 15

Ile Ser Pro Phe Pro Gln Pro Gly Gln Asn Ser Ser Met Val Ser Asn
20 25 30

Ser Ala Gly Ser Asn Ser Arg Gln Leu Pro Leu Pro Leu Ser Ala Asp
35 40 45

Ala Pro Pro Ala Ser Ser Ser His Trp Ser Trp Gln Pro Ser Arg His
50 55 60

Thr Asn Gln Pro Ile Asp Arg Ala Ile Leu Arg Ser Arg Pro Cys Cys
 65 70 75 80

Arg Leu Ser Arg Thr Cys His Trp Ser Leu Gln Pro Pro His Pro Pro
 85 90 95

Arg Arg Ala Ser Gly Trp Ala Ala Ser Leu Gly Arg Ala Ala Val Pro
 100 105 110

Ala Arg Ala Arg Leu Ala Ser Leu Val Arg Ala Gly Ser Ala Gly Arg
 115 120 125

Ala Arg Leu Arg Gly Ala Ala Leu Glu Glu Thr Glu Ala Ala Gly Gly
 130 135 140

Pro Glu Ala Gln Glu Glu Asp Glu Asp Glu Glu Ala Leu Pro His
 145 150 155 160

Ser Glu Ala Met Asp Val Phe Gln Glu Gly Leu Ala Met Val Val Gln
 165 170 175

Asp Pro Leu Leu Cys Asp Leu Pro Ile Gln Val Thr Leu Glu Glu Val
 180 185 190

Asn Ser Gln Ile Ala Leu Glu Tyr Gly Gln Ala Met Thr Val Arg Val
 195 200 205

Cys Lys Met Asp Gly Glu Val Met Arg Lys Cys Tyr Pro Pro Pro Phe
 210 215 220

Arg Leu Cys Gly Pro Gly Phe His Ser Arg Lys Thr
 225 230 235

<210> 179
<211> 143
<212> PRT
<213> Homo sapien

<400> 179

Met Pro Ala Tyr Thr Ala Thr Ala Gly Thr Leu Arg Asp Thr Gln Leu
 1 5 10 15

His Thr His Ile Ala Val His Asn Pro Thr Tyr Asn Gln Lys Thr Lys
 20 25 30

223

His Glu Thr Phe Pro Trp Ala Leu Asn Pro His Val Asn Val His Thr
35 40 45

Gln Thr His Ala Leu Leu Ser His Phe Leu Phe His Thr Pro Ser Ser
50 55 60

Arg Pro Pro Thr Pro Asp Phe Arg His Pro Gln Ser Gln Ser Glu Leu
65 70 75 80

Ala Pro Ala Gln Pro Ser Leu Asp Thr His Ala Pro Pro Thr His Ala
85 90 95

Leu Pro Ser Pro Ala Gly Gly Gly Phe Gly Arg Glu Pro Ala Glu
100 105 110

Pro Ala Ser Asp Ser Arg Cys Gly Ser Asp Ser Ala Leu His Val Leu
115 120 125

Gln Ala Ala Thr Val Ser Glu Ala Arg Arg Gly Arg Glu Leu Glu
130 135 140

<210> 180

<211> 126

<212> PRT

<213> Homo sapien

<400> 180

Ala His Phe Gly Ser Arg Pro Leu Pro Leu Ser Arg Lys Leu Leu Gln
1 5 10 15

Glu Arg His Thr Arg Ser Leu Pro Gln His Cys Lys His Ala Pro Pro
20 25 30

Gln Thr Thr Asn Ala Pro Pro His Thr Arg Leu Leu Ser Leu Thr Lys
35 40 45

Met Pro Ala Tyr Thr Ala Thr Ala Gly Thr Leu Arg Asp Thr Gln Leu
50 55 60

His Thr His Ile Ala Val His Asn Pro Thr Tyr Asn Gln Lys Thr Lys
65 70 75 80

His Glu Thr Phe Pro Trp Ala Leu Asn Pro His Val Asn Val His Thr
85 90 95

Gln Thr His Ala Leu Leu Ser His Phe Leu Phe His Thr Pro Ser Ser
100 105 110

Arg Pro Pro Thr Pro Asp Phe Arg His Pro Gln Ser Gln Ser
115 120 125

<210> 181
<211> 116
<212> PRT
<213> Homo sapien

<400> 181

Ser Ser Ser Ala Cys His Pro Gly Ser Ser Gly Gly Gly Ile Ala Leu
1 5 10 15

Lys Ile Cys Pro Ile Val Lys Gln Glu His Trp Asn Leu His Ser Thr
20 25 30

Ile Arg Pro Cys His Arg Arg Thr Lys Lys Glu Gly Arg Gly Asp His
35 40 45

Ala Pro Ala Ser Arg Glu Ser Pro Phe Phe Ser Ala Ser Tyr Leu Gly
50 55 60

Lys Tyr Lys Gly Val Arg Ala Gly Thr Thr Ser Gln Arg Val His Gly
65 70 75 80

Gly Ser Gly Arg Gly Arg Trp Val Leu His Gly Ala Thr Pro Gly Thr
85 90 95

Phe Leu Leu Ser His Ser Leu Thr Ile Thr Ser Ser Cys Ser Gln Ser
100 105 110

His Ser His Gln
115

<210> 182
<211> 77
<212> PRT
<213> Homo sapien

<400> 182

Lys Pro His Ser Leu Arg Lys Pro Ser Ser Lys Ala Asn Ile Leu Val
1 5 10 15

Ile Cys Glu Lys Ile Glu His Ser Val Ser Leu Leu Leu Ser Ala Ser
20 25 30

Gln His Leu Leu Glu Gln His Glu Leu Leu Thr Leu Thr His Lys Ser

225

35

40

45

Pro Thr Leu Ile Ser Pro Thr Gly Glu Phe Gly Gly Leu Tyr Cys His
50 55 60

Val Pro Gly Ile Ile Cys Ser Ser Leu Tyr Glu Glu
65 70 75

<210> 183
<211> 115
<212> PRT
<213> Homo sapien

<400> 183

Leu Val Phe His Phe Leu Ser Glu Thr Leu Asp Asn Ile Phe Ile Phe
1 5 10 15

Tyr Leu Val Ser Ile Phe Gln Phe Ser Ser Lys Phe Val His Phe Ala
20 25 30

Leu Ser Phe Leu Phe Pro Ser Leu Ser Phe Phe Phe Cys Phe Leu Leu
35 40 45

Phe Arg Phe Lys Phe Ile Phe Phe Leu Leu Lys Val Cys Phe Tyr Leu
50 55 60

Leu Ile Ser Leu Ser Ser Leu Phe Phe Ser Ser Pro Ser Arg Thr Ser
65 70 75 80

Val Phe Gln Phe Ser Thr Ser Asn Phe Tyr Leu Leu Gln Ile Val Ser
85 90 95

Ser Tyr His Ser Gln Leu Ile Phe Pro Phe Ser Ser Ala Phe Ser Lys
100 105 110

Cys Val Asn
115

<210> 184
<211> 84
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (77)..(78)
<223> X=any amino acid

226

<220>
<221> MISC_FEATURE
<222> (82)...(82)
<223> X=any amino acid

<400> 184

Lys Pro His Ser Leu Arg Lys Pro Ser Ser Lys Ala Asn Ile Leu Val
1 5 10 15

Ile Cys Glu Lys Ile Glu His Ser Val Ser Leu Leu Leu Ser Ala Ser
20 25 30

Gln His Leu Leu Glu Gln His Glu Leu Leu Thr Leu Thr His Lys Ser
35 40 45

Pro Thr Leu Ile Ser Pro Thr Gly Glu Phe Gly Gly Leu Tyr Cys His
50 55 60

Val Pro Gly Ile Ile Cys Ser Ser Leu Tyr Glu Xaa Xaa Asn Leu
65 70 75 80

Ser Xaa Leu Pro

<210> 185
<211> 84
<212> PRT
<213> Homo sapien

<220>
<221> MISC_FEATURE
<222> (77)...(78)
<223> X=any amino acid

<220>
<221> MISC_FEATURE
<222> (82)...(82)
<223> X=any amino acid

<400> 185

Lys Pro His Ser Leu Arg Lys Pro Ser Ser Lys Ala Asn Ile Leu Val
1 5 10 15

Ile Cys Glu Lys Ile Glu His Ser Val Ser Leu Leu Leu Ser Ala Ser
20 25 30

Gln His Leu Leu Glu Gln His Glu Leu Leu Thr Leu Thr His Lys Ser
35 40 45

Pro Thr Leu Ile Ser Pro Thr Gly Glu Phe Gly Gly Leu Tyr Cys His
 50 55 60

Val Pro Gly Ile Ile Cys Ser Ser Leu Tyr Glu Xaa Xaa Asn Leu
 65 70 75 80

Ser Xaa Leu Pro

<210> 186
 <211> 104
 <212> PRT
 <213> Homo sapien

<400> 186

Met Val Leu Cys Lys Ile Lys Gln His Val Glu Gly Ile Val Ser Ala
 1 5 10 15

Trp Trp Leu Leu Glu Pro Pro Glu Arg Cys Cys Gly Ser Ser Thr Ser
 20 25 30

Ala Thr Asn Ser Thr Ser Val Ser Ser Arg Lys Ala Glu Asn Lys Tyr
 35 40 45

Ala Gly Gly Asn Pro Val Cys Val Arg Pro Thr Pro Lys Trp Gln Lys
 50 55 60

Gly Ile Gly Glu Phe Phe Arg Leu Ser Pro Lys Asp Ser Glu Lys Glu
 65 70 75 80

Asn Gln Ile Pro Glu Glu Ala Gly Ser Ser Gly Leu Gly Lys Ala Lys
 85 90 95

Arg Lys Ala Cys Pro Cys Ala Thr
 100

<210> 187
 <211> 107
 <212> PRT
 <213> Homo sapien

<400> 187

Asn Lys Thr Ala Arg Gly Arg Tyr Cys Lys Arg Leu Val Ala Ala Arg
 1 5 10 15

Ala Pro Arg Lys Val Leu Gly Ser Ser Thr Ser Ala Thr Asn Ser Thr

228

20

25

30

Ser Val Ser Ser Arg Lys Ala Glu Asn Lys Tyr Ala Gly Gly Asn Pro
35 40 45

Val Cys Val Arg Pro Thr Pro Lys Trp Gln Lys Gly Ile Gly Glu Phe
50 55 60

Phe Arg Leu Ser Pro Lys Asp Ser Glu Lys Glu Asn Gln Ile Pro Glu
65 70 75 80

Glu Ala Gly Ser Ser Gly Leu Gly Lys Ala Lys Arg Lys Ala Cys Pro
85 90 95

Leu Gln Pro Asp His Thr Asn Asp Glu Lys Glu
100 105

<210> 188

<211> 38

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (12)..(12)

<223> X=any amino acid

<400> 188

Pro Pro Pro Arg Leu Leu Ile Tyr Lys Gly Gln Xaa Val Ile Leu Asp
1 5 10 15

Ala Ala Arg Ala Ala Gln Cys Asp Gly Leu Val Ala Ala Glu Val Pro
20 25 30

Asp Tyr Asn Ala Arg Ile
35

<210> 189

<211> 47

<212> PRT

<213> Homo sapien

<400> 189

Ile Phe Val Leu Ile Asn Leu Val Asn Lys Asn Lys Ser Lys Ser Glu
1 5 10 15

Lys Lys Thr Thr Gln Lys Lys Val Gly Gly Asn Gln Gly Pro Lys
20 25 30

Gly Ser Leu Cys Asp Leu Val Phe Arg Pro Ile Pro Gln Val Gly
35 40 45

<210> 190
<211> 71
<212> PRT
<213> Homo sapien

<400> 190

Met Leu Leu Glu Arg Arg Ser Val Asp Gly Ser Trp Ser Arg Pro Arg
1 5 10 15

Tyr Ile Asp Phe Thr Ala Asp Gln Val Asp Leu Thr Ser Ala Leu Thr
20 25 30

Lys Lys Ile Thr Leu Lys Thr Pro Leu Val Ser Ser Pro Met Asp Thr
35 40 45

Val Thr Glu Ala Gly Met Ala Ile Ala Met Ala Leu Thr Gly Gly Ile
50 55 60

Gly Phe Ile His His Asn Ser
65 70

<210> 191
<211> 138
<212> PRT
<213> Homo sapien

<400> 191

Met Pro Ile Thr Ser Thr Ser Pro Val Glu Pro Val Val Thr Thr Glu
1 5 10 15

Gly Ser Ser Gly Ala Ala Gly Leu Glu Pro Arg Lys Leu Ser Ser Lys
20 25 30

Thr Arg Arg Asp Lys Glu Lys Gln Ser Cys Lys Ser Cys Gly Glu Thr
35 40 45

Phe Asn Ser Ile Thr Lys Arg Arg His His Cys Lys Leu Cys Gly Ala
50 55 60

Val Ile Cys Gly Lys Cys Ser Glu Phe Lys Ala Glu Asn Ser Arg Gln
65 70 75 80

Ser Arg Val Cys Arg Asp Cys Phe Leu Thr Gln Pro Val Ala Pro Glu

230

85

90

95

Ser Thr Glu Val Gly Ala Pro Ser Ser Cys Ser Pro Pro Gly Gly Ala
100 105 110

Ala Glu Pro Pro Asp Thr Cys Ser Cys Ala Pro Ala Ala Leu Ala Ala
115 120 125

Ser Ala Phe Gly Val Ser Leu Gly Pro Gly
130 135

<210> 192

<211> 67

<212> PRT

<213> Homo sapien

<400> 192

Ser Arg Gly Ser Arg Leu Pro Ser Asn Phe Pro Ser Asp Leu Tyr Ser
1 5 10 15

Leu Ala His Ser Tyr Leu Gly Gly Gly Arg Lys Gly Arg Thr Lys
20 25 30

Arg Glu Ala Ala Ala Asn Thr Asn Arg Pro Ser Pro Gly Gly His Glu
35 40 45

Arg Lys Leu Val Thr Lys Leu Gln Asn Ser Glu Arg Lys Lys Arg Gly
50 55 60

Ala Arg Arg

65

<210> 193

<211> 65

<212> PRT

<213> Homo sapien

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X=any amino acid

<220>

<221> MISC_FEATURE

<222> (13)..(13)

<223> X=any amino acid

<400> 193

231

Leu Glu Asp Leu Gly Cys Leu Ala Leu Xaa Ser Asp Xaa Ile Ala Gly
1 5 10 15

His Ser Tyr Leu Gly Gly Gly Arg Lys Gly Arg Thr Lys Arg Glu
20 25 30

Ala Ala Ala Asn Thr Asn Arg Pro Ser Pro Gly Gly His Glu Arg Lys
35 40 45

Leu Val Thr Lys Leu Gln Asn Ser Glu Arg Lys Lys Arg Gly Ala Arg
50 55 60

Arg
65

<210> 194
<211> 195
<212> PRT
<213> Homo sapien

<400> 194

Met Gly Ser His Tyr Val Ser Gln Ala Asp Pro Lys Phe Leu Gly Ser
1 5 10 15

Ser Asn Ser Pro Ala Leu Ala Ser Gln Ser Ala Glu Ile Thr Gly Val
20 25 30

Ser His Pro Ala Gln Pro Thr His Pro Phe Leu Ala Asn Leu Phe Leu
35 40 45

Gly Pro Ser Arg His Pro Cys Leu Ile Pro Tyr Pro Arg Ser Ala Met
50 55 60

Leu Leu Ser Leu Gly Pro His Thr His Leu Gly Ser His Ile Pro Gln
65 70 75 80

Arg Gly Ser Ser Arg Leu Leu Pro Ala Leu Pro Ile Pro Thr Thr Leu
85 90 95

Asn Pro Cys Leu Ser Ser Asp Arg Ala Ser His His Ala Tyr Ala His
100 105 110

Phe Thr Ser Asp Ser Cys Leu Gly Tyr Arg Arg Trp Arg Pro Glu Arg
115 120 125

Ser His Gln Glu Arg Ser Cys Cys Gln His Gln Pro Pro Gln Pro Trp
130 135 140

Arg Ala Arg Glu Glu Thr Gly Asp Gln Ala Ala Glu Phe Arg Glu Glu
145 150 155 160

Glu Ala Arg Gly Thr Ala Leu Arg Gln Ser Trp Arg Val Arg Ser Arg
165 170 175

Gly Ala Gln Arg Ala Gln Gly Gly Ala Ser Ala Met Lys Asp Arg Pro
180 185 190

Glu Gly Val
195

<210> 195
<211> 124
<212> PRT
<213> Homo sapien

<400> 195

Trp Met Trp Ser Arg Pro Arg Trp Gly Ala Glu Phe Arg Lys Ile Pro
1 5 10 15

Thr Ser Met Lys Ala Lys Arg Ser His Gln Ala Ile Ile Met Ser Thr
20 25 30

Ser Leu Arg Val Ser Pro Ser Ile His Gly Tyr His Phe Asp Thr Ala
35 40 45

Ser Arg Lys Lys Ala Val Gly Asn Ile Phe Glu Asn Thr Asp Gln Glu
50 55 60

Ser Leu Glu Arg Leu Phe Arg Asn Ser Gly Asp Lys Lys Ala Glu Glu
65 70 75 80

Arg Ala Lys Ile Ile Phe Ala Ile Asp Gln Asp Val Glu Glu Lys Thr
85 90 95

Arg Ala Leu Met Ala Leu Lys Lys Arg Thr Lys Asp Lys Leu Phe Gln
100 105 110

Phe Leu Lys Leu Arg Lys Tyr Ser Ile Lys Val His
115 120

<210> 196
<211> 106
<212> PRT
<213> Homo sapien

<400> 196

Met Lys Ala Lys Arg Ser His Gln Ala Ile Ile Met Ser Thr Ser Leu
1 5 10 15

Arg Val Ser Pro Ser Ile His Gly Tyr His Phe Asp Thr Ala Ser Arg
20 25 30

Lys Lys Ala Val Gly Asn Ile Phe Glu Asn Thr Asp Gln Glu Ser Leu
35 40 45

Glu Arg Leu Phe Arg Asn Ser Gly Asp Lys Lys Ala Glu Glu Arg Ala
50 55 60

Lys Ile Ile Phe Ala Ile Asp Gln Asp Val Glu Glu Lys Thr Arg Ala
65 70 75 80

Leu Met Ala Leu Lys Lys Arg Thr Lys Asp Lys Leu Phe Gln Phe Leu
85 90 95

Lys Leu Arg Lys Tyr Ser Ile Lys Val His
100 105

<210> 197

<211> 129

<212> PRT

<213> Homo sapien

<400> 197

Met Leu Leu Glu Arg Arg Ser Val Met Asp Gly Gln Val Lys Gly Ala
1 5 10 15

Glu Phe Arg Lys Ile Pro Thr Ser Met Lys Ala Lys Arg Ser His Gln
20 25 30

Ala Ile Ile Met Ser Thr Ser Leu Arg Val Ser Pro Ser Ile His Gly
35 40 45

Tyr His Phe Asp Thr Ala Ser Arg Lys Lys Ala Val Gly Asn Ile Phe
50 55 60

Glu Asn Thr Asp Gln Glu Ser Leu Glu Arg Leu Phe Arg Asn Ser Gly
65 70 75 80

Asp Lys Lys Ala Glu Glu Arg Ala Lys Ile Ile Phe Ala Ile Asp Gln
85 90 95

234

Asp Val Glu Glu Lys Thr Arg Ala Leu Met Ala Leu Lys Lys' Arg Thr
 100 105 110

Lys Cys Phe Gln Gln Gly Phe Glu Asn Ser Ser Val Pro Ala Gly Lys
 115 120 125

Asp

<210> 198
<211> 130
<212> PRT
<213> Homo sapien

<400> 198

Met Leu Leu Glu Arg Arg Ser Val Met Asp Gly Gln Val Ser Leu Gly
 1 5 10 15

Ala Glu Phe Arg Lys Ile Pro Thr Ser Met Lys Ala Lys Arg Ser His
 20 25 30

Gln Ala Ile Ile Met Ser Thr Ser Leu Arg Val Ser Pro Ser Ile His
 35 40 45

Gly Tyr His Phe Asp Thr Ala Ser Arg Lys Lys Ala Val Gly Asn Ile
 50 55 60

Phe Glu Asn Thr Asp Gln Glu Ser Leu Glu Arg Leu Phe Arg Asn Ser
 65 70 75 80

Gly Asp Lys Lys Ala Glu Glu Arg Ala Lys Ile Ile Phe Ala Ile Asp
 85 90 95

Gln Asp Val Glu Glu Lys Thr Arg Ala Leu Met Ala Leu Lys Lys Arg
 100 105 110

Thr Lys Cys Phe Gln Gln Gly Phe Glu Asn Ser Ser Val Pro Ala Gly
 115 120 125

Lys Asp
 130

<210> 199
<211> 85
<212> PRT
<213> Homo sapien

235

<400> 199

Ile Leu Cys Asp Met Ile Phe Trp Ile Tyr Arg Thr Leu Ala His Val
1 5 10 15

Pro Cys Ala Ser His Ser Ser Glu Val Ile Ile Tyr Thr Glu Gly Phe
20 25 30

Lys Ile Arg Leu Glu Val Glu Ile Tyr Tyr Leu Phe Met His Cys Thr
35 40 45

Val Phe Leu Tyr Cys Cys Leu Lys Leu Leu Ser Cys Ala Ser Leu Ile
50 55 60

Lys Ala Gln Asn Val Leu Pro Thr Pro Tyr Leu Arg Arg Asn Lys Ile
65 70 75 80

Thr Ser Ile Asp Phe
85

<210> 200

<211> 68

<212> PRT

<213> Homo sapien

<400> 200

Asp Ala Cys Arg Ala Gly Arg Ser Val Asp Gly Tyr Lys Ala Val Arg
1 5 10 15

Phe Ser Ser Pro Ser Arg Ala Leu Leu Gly Thr Arg Glu Ile Trp Leu
20 25 30

Trp Ser Arg Trp Ser Ser Leu Thr Pro His Arg Ala Asn Leu Asn Leu
35 40 45

Val Leu Glu Lys Ala Phe Ser Asn Ser Thr Pro Pro Tyr Lys Met His
50 55 60

Met Glu Val Gly
65

<210> 201

<211> 378

<212> PRT

<213> Homo sapien

<400> 201

Ser Ala Val Gly Ser Asp His Ile Phe His Asn Ile Pro Gly Ser Thr

236

1	5	10	15
---	---	----	----

Ser Ser Ala Thr Asn Val Ser Met Val Val Ser Ala Gly Pro Trp Ser
20 25 30

Ser Glu Lys Ala Glu Thr Asn Ile Leu Glu Ile Asn Glu Lys Leu Arg
35 40 45

Pro Gln Leu Ala Glu Asn Lys Gln Gln Phe Arg Asn Leu Lys Glu Lys
50 55 60

Cys Phe Val Thr Gln Leu Ala Gly Phe Leu Ala Asn Arg Gln Lys Lys
65 70 75 80

Tyr Lys Tyr Glu Glu Cys Lys Asp Leu Ile Lys Phe Met Leu Arg Asn
85 90 95

Glu Arg Gln Phe Lys Glu Glu Lys Leu Ala Glu Gln Leu Lys Gln Ala
100 105 110

Glu Glu Leu Arg Gln Tyr Lys Val Leu Val His Ser Gln Glu Arg Glu
115 120 125

Leu Thr Gln Leu Arg Glu Lys Leu Arg Glu Gly Arg Asp Ala Ser Arg
130 135 140

Ser Leu Asn Gln His Leu Gln Ala Leu Leu Thr Pro Asp Glu Pro Asp
145 150 155 160

Lys Ser Gln Gly Gln Asp Leu Gln Glu Gln Leu Ala Glu Gly Cys Arg
165 170 175

Leu Ala Gln His Leu Val Gln Lys Leu Ser Pro Glu Asn Asp Asn Asp
180 185 190

Asp Asp Glu Asp Val Gln Val Glu Val Ala Glu Lys Val Gln Lys Ser
195 200 205

Ser Ala Pro Arg Glu Met Pro Lys Ala Glu Glu Lys Glu Val Pro Glu
210 215 220

Asp Ser Leu Glu Glu Cys Ala Ile Thr Cys Ser Asn Ser His Gly Pro
225 230 235 240

Tyr Asp Ser Asn Gln Pro His Arg Lys Thr Lys Ile Thr Phe Glu Glu
245 250 255

Asp Lys Val Asp Ser Thr Leu Ile Gly Ser Ser Ser His Val Glu Trp
260 265 270

Glu Asp Ala Val His Ile Ile Pro Glu Asn Glu Ser Asp Asp Glu Glu
275 280 285

Glu Glu Glu Lys Gly Pro Val Ser Pro Arg Asn Leu Gln Glu Ser Glu
290 295 300

Glu Glu Glu Val Pro Gln Glu Ser Trp Asp Glu Gly Tyr Ser Thr Leu
305 310 315 320

Ser Ile Pro Pro Glu Met Leu Ala Ser Tyr Gln Ser Tyr Ser Gly Thr
325 330 335

Phe His Ser Leu Glu Glu Gln Gln Val Cys Met Ala Val Asp Ile Gly
340 345 350

Gly His Arg Trp Asp Gln Val Lys Lys Glu Asp Gln Glu Ala Thr Gly
355 360 365

Pro Ser Gln Ala Gln Gln Gly Ala Ala Gly
370 375

<210> 202
<211> 876
<212> PRT
<213> Homo sapien

<400> 202

Met Gly Asn Ser Lys Lys Asn Thr Glu Thr Gly Lys Thr Thr Phe Phe
1 5 10 15

Thr Asn Glu Leu Phe Ile His Phe Gln Trp Ile Gln Thr Lys Leu Gln
20 25 30

Lys Thr Gln Arg Lys Ser Gly Gln Ala Lys Ser Leu Ile Ser Tyr Thr
35 40 45

Cys Gly Lys Ala Leu Ser Ser Val Leu Thr Glu Ser Arg Trp Gly Asp
50 55 60

Phe Met Thr Thr Ile Lys Lys Ile Gln Leu Leu Gly Asn Cys Phe Cys
65 70 75 80

238

Leu Asp Asp Val Val Gln Thr Arg Asp Lys Gln Leu Arg Asn Met Leu
85 90 95

Arg Cys Ile Gly Lys Asp Thr Gly Leu Trp His His His Lys Gly Thr
100 105 110

Arg Ile Leu Arg Val Asn Ala Glu Gly Met Ile Pro Ile Gly Gly Asp
115 120 125

Pro Gln Val Arg Leu Gly Cys Leu Cys Phe Arg Lys Ala Trp Ala Ile
130 135 140

Gly Met Gln Gly Ser Tyr Asp Ser Met Thr Pro Pro Pro Ser Asn Ser
145 150 155 160

Val Ile Ala Thr Ala Asp Gly Tyr Leu Ala Arg Trp Pro Gln Ser Thr
165 170 175

Ser Leu Leu Ser Glu Ser Glu Leu Leu Ala Val Leu Ser Ala Leu Ser
180 185 190

Ser Gly Thr Ser Asn Leu Val Phe Val Val Lys Asp Pro Lys Val Leu
195 200 205

Trp Gly Val Ile Thr Phe Phe Tyr Asn Ile Pro Gly Ser Thr Ser Ser
210 215 220

Ala Thr Asn Val Ser Met Val Val Ser Ala Gly Pro Trp Ser Ser Glu
225 230 235 240

Lys Ala Glu Thr Asn Ile Leu Glu Ile Asn Glu Lys Leu Arg Pro Gln
245 250 255

Leu Ala Glu Asn Lys Gln Gln Phe Arg Asn Leu Lys Glu Lys Cys Phe
260 265 270

Val Thr Gln Leu Ala Gly Phe Leu Ala Asn Arg Gln Lys Lys Tyr Lys
275 280 285

Tyr Glu Glu Cys Lys Asp Leu Ile Lys Phe Met Leu Arg Asn Glu Arg
290 295 300

Gln Phe Lys Glu Glu Lys Leu Ala Glu Gln Leu Lys Gln Ala Glu Glu
305 310 315 320

Leu Arg Gln Tyr Lys Val Leu Val His Ser Gln Glu Arg Glu Leu Thr

325	330	335
Gln Leu Arg Glu Lys Leu Arg Glu Gly Arg Asp Ala Ser Cys Ser Leu 340	345	350
Asn Gln His Leu Gln Ala Leu Leu Thr Pro Asp Glu Pro Asp Lys Ser 355	360	365
Gln Gly Gln Asp Leu Gln Glu Gln Leu Ala Glu Gly Cys Arg Leu Ala 370	375	380
Gln His Leu Val Gln Lys Leu Ser Pro Glu Asn Asp Asn Asp Asp Asp 385	390	395
Glu Asp Val Gln Val Glu Val Ala Glu Lys Val Gln Lys Ser Ser Ala 405	410	415
Pro Arg Glu Met Pro Lys Ala Glu Glu Lys Glu Val Pro Glu Asp Ser 420	425	430
Leu Glu Glu Cys Ala Ile Thr Cys Ser Asn Ser His Gly Pro Tyr Asp 435	440	445
Ser Asn Gln Pro His Arg Lys Thr Lys Ile Thr Phe Glu Glu Asp Lys 450	455	460
Val Asp Ser Thr Leu Ile Gly Ser Ser Ser His Val Glu Trp Glu Asp 465	470	475
Ala Val His Ile Ile Pro Glu Asn Glu Ser Asp Asp Glu Glu Glu 485	490	495
Glu Lys Gly Pro Val Ser Pro Arg Asn Leu Gln Glu Ser Glu Glu Glu 500	505	510
Glu Val Pro Gln Glu Ser Trp Asp Glu Gly Tyr Ser Thr Leu Ser Ile 515	520	525
Pro Pro Glu Met Leu Ala Ser Tyr Gln Ser Tyr Ser Gly Thr Phe His 530	535	540
Ser Leu Glu Glu Gln Gln Val Cys Met Ala Val Asp Ile Gly Gly His 545	550	555
Arg Trp Asp Gln Val Lys Lys Glu Asp Gln Glu Ala Thr Gly Pro Ser 565	570	575

Gln Leu Ser Arg Glu Leu Leu Asp Glu Lys Gly Pro Glu Val Leu Gln
580 585 590

Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser Gly Tyr Leu Glu Leu
595 600 605

Thr Asp Ser Cys Gln Pro Tyr Arg Ser Ala Phe Tyr Ile Leu Glu Gln
610 615 620

Gln Arg Val Gly Trp Ala Leu Asp Met Asp Glu Ile Glu Lys Tyr Gln
625 630 635 640

Glu Val Glu Glu Asp Gln Asp Pro Ser Cys Pro Arg Leu Ser Arg Glu
645 650 655

Leu Leu Asp Glu Lys Glu Pro Glu Val Leu Gln Asp Ser Leu Asp Arg
660 665 670

Cys Tyr Ser Thr Pro Ser Gly Tyr Leu Glu Leu Pro Asp Leu Gly Gln
675 680 685

Pro Tyr Arg Ser Ala Val His Ser Leu Glu Glu Gln Tyr Leu Gly Leu
690 695 700

Ala Leu Asp Val Asp Arg Ile Lys Lys Asp Gln Glu Glu Glu Asp
705 710 715 720

Gln Gly Pro Pro Cys Pro Arg Leu Ser Arg Glu Leu Leu Glu Ala Val
725 730 735

Glu Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro
740 745 750

Ser Ser Cys Leu Glu Gln Pro Asp Ser Cys Leu Pro Tyr Gly Ser Ser
755 760 765

Phe Tyr Ala Leu Glu Glu Lys His Val Gly Phe Ser Leu Asp Val Gly
770 775 780

Glu Ile Glu Lys Lys Gly Lys Lys Lys Arg Arg Gly Arg Arg Ser
785 790 795 800

Thr Lys Lys Arg Arg Arg Gly Arg Lys Glu Gly Glu Glu Asp Gln
805 810 815

241

Asn Pro Pro Cys Pro Arg Leu Ser Arg Glu Leu Leu Asp Glu Lys Gly
 820 825 830

Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser
 835 840 845

Gly Tyr Leu Glu Leu Thr Asp Ser Cys Gln Pro Tyr Arg Ser Ala Phe
 850 855 860

Tyr Leu Leu Glu Gln Gln Arg Val Glu Leu Arg Pro
 865 870 875

<210> 203
<211> 378
<212> PRT
<213> Homo sapien

<400> 203

Ser Ala Val Gly Ser Asp His Ile Phe His Asn Ile Pro Gly Ser Thr
 1 5 10 15

Ser Ser Ala Thr Asn Val Ser Met Val Val Ser Ala Gly Pro Trp Ser
 20 25 30

Ser Glu Lys Ala Glu Thr Asn Ile Leu Glu Ile Asn Glu Lys Leu Arg
 35 40 45

Pro Gln Leu Ala Glu Asn Lys Gln Gln Phe Arg Asn Leu Lys Glu Lys
 50 55 60

Cys Phe Val Thr Gln Leu Ala Gly Phe Leu Ala Asn Arg Gln Lys Lys
 65 70 75 80

Tyr Lys Tyr Glu Glu Cys Lys Asp Leu Ile Lys Phe Met Leu Arg Asn
 85 90 95

Glu Arg Gln Phe Lys Glu Glu Lys Leu Ala Glu Gln Leu Lys Gln Ala
 100 105 110

Glu Glu Leu Arg Gln Tyr Lys Val Leu Val His Ser Gln Glu Arg Glu
 115 120 125

Leu Thr Gln Leu Arg Glu Lys Leu Arg Glu Gly Arg Asp Ala Ser Arg
 130 135 140

Ser Leu Asn Gln His Leu Gln Ala Leu Leu Thr Pro Asp Glu Pro Asp

242

145

150

155

160

Lys Ser Gln Gly Gln Asp Leu Gln Glu Gln Leu Ala Glu Gly Cys Arg
165 170 175

Leu Ala Gln His Leu Val Gln Lys Leu Ser Pro Glu Asn Asn Asn Asp
180 185 190

Asp Asp Glu Asp Val Gln Val Glu Val Ala Glu Lys Val Gln Lys Ser
195 200 205

Ser Ala Pro Arg Glu Met Pro Lys Ala Glu Glu Lys Glu Val Pro Glu
210 215 220

Asp Ser Leu Glu Glu Cys Ala Ile Thr Cys Ser Asn Ser His Gly Pro
225 230 235 240

Tyr Asp Ser Asn Gln Pro His Arg Lys Thr Lys Ile Thr Phe Glu Glu
245 250 255

Asp Lys Val Asp Ser Thr Leu Ile Gly Ser Ser Ser His Val Glu Trp
260 265 270

Glu Asp Ala Val His Ile Ile Pro Glu Asn Glu Ser Asp Asp Glu Glu
275 280 285

Glu Glu Glu Lys Gly Pro Val Ser Pro Arg Asn Leu Gln Glu Ser Glu
290 295 300

Glu Glu Glu Val Pro Gln Glu Ser Trp Asp Glu Gly Tyr Ser Thr Leu
305 310 315 320

Ser Ile Pro Pro Glu Met Leu Ala Ser Tyr Gln Ser Tyr Ser Gly Thr
325 330 335

Phe His Ser Leu Glu Glu Gln Gln Val Cys Met Ala Val Asp Ile Gly
340 345 350

Gly His Arg Trp Asp Gln Val Lys Lys Glu Asp Gln Glu Ala Thr Gly
355 360 365

Pro Ser Gln Ala Gln Gln Gly Ala Ala Gly
370 375

<210> 204

<211> 782

243

<212> PRT
<213> Homo sapien

<400> 204

Met Leu Arg Cys Ile Gly Lys Asp Thr Gly Leu Trp His His His Lys
1 5 10 15

Gly Thr Arg Ile Leu Arg Val Asn Ala Glu Gly Met Ile Pro Ile Gly
20 25 30

Gly Asp Pro Gln Val Arg Leu Gly Cys Leu Cys Phe Arg Lys Ala Trp
35 40 45

Ala Ile Gly Met Gln Gly Ser Tyr Asp Ser Met Thr Pro Pro Pro Ser
50 55 60

Asn Ser Val Ile Ala Thr Ala Asp Gly Tyr Leu Ala Arg Trp Pro Gln
65 70 75 80

Ser Thr Ser Leu Leu Ser Glu Ser Glu Leu Leu Ala Val Leu Ser Ala
85 90 95

Leu Ser Ser Gly Thr Ser Asn Leu Val Phe Val Val Lys Asp Pro Lys
100 105 110

Val Leu Trp Gly Val Ile Thr Phe Phe Tyr Asn Ile Pro Gly Ser Thr
115 120 125

Ser Ser Ala Thr Asn Val Ser Met Val Val Ser Ala Gly Pro Trp Ser
130 135 140

Ser Glu Lys Ala Glu Thr Asn Ile Leu Glu Ile Asn Glu Lys Leu Arg
145 150 155 160

Pro Gln Leu Ala Glu Asn Lys Gln Gln Phe Arg Asn Leu Lys Glu Lys
165 170 175

Cys Phe Val Thr Gln Leu Ala Gly Phe Leu Ala Asn Arg Gln Lys Lys
180 185 190

Tyr Lys Tyr Glu Glu Cys Lys Asp Leu Ile Lys Phe Met Leu Arg Asn
195 200 205

Glu Arg Gln Phe Lys Glu Glu Lys Leu Ala Glu Gln Leu Lys Gln Ala
210 215 220

244

Glu Glu Leu Arg Gln Tyr Lys Val Leu Val His Ser Gln Glu Arg Glu
225 230 235 240

Leu Thr Gln Leu Arg Glu Lys Leu Arg Glu Gly Arg Asp Ala Ser Cys
245 250 255

Ser Leu Asn Gln His Leu Gln Ala Leu Leu Thr Pro Asp Glu Pro Asp
260 265 270

Lys Ser Gln Gly Gln Asp Leu Gln Glu Gln Leu Ala Glu Gly Cys Arg
275 280 285

Leu Ala Gln His Leu Val Gln Lys Leu Ser Pro Glu Asn Asp Asn Asp
290 295 300

Asp Asp Glu Asp Val Gln Val Glu Val Ala Glu Lys Val Gln Lys Ser
305 310 315 320

Ser Ala Pro Arg Glu Met Pro Lys Ala Glu Glu Lys Glu Val Pro Glu
325 330 335

Asp Ser Leu Glu Glu Cys Ala Ile Thr Cys Ser Asn Ser His Gly Pro
340 345 350

Tyr Asp Ser Asn Gln Pro His Arg Lys Thr Lys Ile Thr Phe Glu Glu
355 360 365

Asp Lys Val Asp Ser Thr Leu Ile Gly Ser Ser Ser His Val Glu Trp
370 375 380

Glu Asp Ala Val His Ile Ile Pro Glu Asn Glu Ser Asp Asp Glu Glu
385 390 395 400

Glu Glu Glu Lys Gly Pro Val Ser Pro Arg Asn Leu Gln Glu Ser Glu
405 410 415

Glu Glu Glu Val Pro Gln Glu Ser Trp Asp Glu Gly Tyr Ser Thr Leu
420 425 430

Ser Ile Pro Pro Glu Met Leu Ala Ser Tyr Gln Ser Tyr Ser Gly Thr
435 440 445

Phe His Ser Leu Glu Glu Gln Gln Val Cys Met Ala Val Asp Ile Gly
450 455 460

Gly His Arg Trp Asp Gln Val Lys Lys Glu Asp Gln Glu Ala Thr Gly

245

465

470

475

480

Pro Ser Gln Leu Ser Arg Glu Leu Leu Asp Glu Lys Gly Pro Glu Val
485 490 495

Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr Pro Ser Gly Tyr Leu
500 505 510

Glu Leu Thr Asp Ser Cys Gln Pro Tyr Arg Ser Ala Phe Tyr Ile Leu
515 520 525

Glu Gln Gln Arg Val Gly Trp Ala Leu Asp Met Asp Glu Ile Glu Lys
530 535 540

Tyr Gln Glu Val Glu Glu Asp Gln Asp Pro Ser Cys Pro Arg Leu Ser
545 550 555 560

Arg Glu Leu Leu Asp Glu Lys Glu Pro Glu Val Leu Gln Asp Ser Leu
565 570 575

Asp Arg Cys Tyr Ser Thr Pro Ser Gly Tyr Leu Glu Leu Pro Asp Leu
580 585 590

Gly Gln Pro Tyr Arg Ser Ala Val His Ser Leu Glu Glu Gln Tyr Leu
595 600 605

Gly Leu Ala Leu Asp Val Asp Arg Ile Lys Lys Asp Gln Glu Glu Glu
610 615 620

Glu Asp Gln Gly Pro Pro Cys Pro Arg Leu Ser Arg Glu Leu Leu Glu
625 630 635 640

Ala Val Glu Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser
645 650 655

Thr Pro Ser Ser Cys Leu Glu Gln Pro Asp Ser Cys Leu Pro Tyr Gly
660 665 670

Ser Ser Phe Tyr Ala Leu Glu Glu Lys His Val Gly Phe Ser Leu Asp
675 680 685

Val Gly Glu Ile Glu Lys Lys Gly Lys Lys Lys Arg Arg Gly Arg
690 695 700

Arg Ser Thr Lys Lys Arg Arg Arg Arg Gly Arg Lys Glu Gly Glu Glu
705 710 715 720

Asp Gln Asn Pro Pro Cys Pro Arg Leu Ser Arg Glu Leu Leu Asp Glu
725 730 735

Lys Gly Pro Glu Val Leu Gln Asp Ser Leu Asp Arg Cys Tyr Ser Thr
740 745 750

Pro Ser Gly Tyr Leu Glu Leu Thr Asp Ser Cys Gln Pro Tyr Arg Ser
755 760 765

Ala Phe Tyr Leu Leu Glu Gln Gln Arg Val Glu Leu Arg Pro
770 775 780

<210> 205

<211> 449

<212> PRT

<213> Homo sapien

<400> 205

Met Ala Phe Ala Arg Arg Leu Leu Arg Gly Pro Leu Ser Gly Pro Leu
1 5 10 15

Leu Gly Arg Arg Gly Val Cys Ala Gly Ala Met Ala Pro Pro Arg Arg
20 25 30

Phe Val Leu Glu Leu Pro Asp Cys Thr Leu Ala His Phe Ala Leu Gly
35 40 45

Ala Asp Ala Pro Gly Asp Ala Asp Ala Pro Asp Pro Arg Leu Ala Ala
50 55 60

Leu Leu Gly Pro Pro Glu Arg Ser Tyr Ser Leu Cys Val Pro Val Thr
65 70 75 80

Pro Asp Ala Gly Cys Gly Ala Arg Val Arg Ala Ala Arg Leu His Gln
85 90 95

Arg Leu Leu His Gln Leu Arg Arg Gly Pro Phe Gln Arg Cys Gln Leu
100 105 110

Leu Arg Leu Leu Cys Tyr Cys Pro Gly Gly Gln Ala Gly Gly Ala Gln
115 120 125

Gln Gly Phe Leu Leu Arg Asp Pro Leu Asp Asp Pro Asp Thr Arg Gln
130 135 140

247

Ala Leu Leu Glu Leu Leu Gly Ala Cys Gln Glu Ala Pro Arg Pro His
145 150 155 160

Leu Gly Glu Phe Glu Ala Asp Pro Arg Gly Gln Leu Trp Gln Arg Leu
165 170 175

Trp Glu Val Gln Asp Gly Arg Arg Leu Gln Val Gly Cys Ala Gln Val
180 185 190

Val Pro Val Pro Glu Pro Pro Leu His Pro Val Val Pro Asp Leu Pro
195 200 205

Ser Ser Val Val Phe Pro Asp Arg Glu Ala Ala Arg Ala Val Leu Glu
210 215 220

Glu Cys Thr Ser Phe Ile Pro Glu Ala Arg Ala Val Leu Asp Leu Val
225 230 235 240

Asp Gln Cys Pro Lys Gln Ile Gln Lys Gly Lys Phe Gln Val Val Ala
245 250 255

Ile Glu Gly Leu Asp Ala Thr Gly Lys Thr Thr Val Thr Gln Ser Val
260 265 270

Ala Asp Ser Leu Lys Ala Val Leu Leu Lys Ser Pro Pro Ser Cys Ile
275 280 285

Gly Gln Trp Arg Lys Ile Phe Asp Asp Glu Pro Thr Ile Ile Arg Arg
290 295 300

Ala Phe Tyr Ser Leu Gly Asn Tyr Ile Val Ala Ser Glu Ile Ala Lys
305 310 315 320

Glu Ser Ala Lys Ser Pro Val Ile Val Asp Arg Tyr Trp His Ser Thr
325 330 335

Ala Thr Tyr Ala Ile Ala Thr Glu Val Ser Gly Gly Leu Gln His Leu
340 345 350

Pro Pro Ala His His Pro Val Tyr Gln Trp Pro Glu Asp Leu Leu Lys
355 360 365

Pro Asp Leu Ile Leu Leu Leu Thr Val Ser Pro Glu Glu Arg Leu Gln
370 375 380

Arg Leu Gln Gly Arg Gly Met Glu Lys Thr Arg Glu Glu Ala Glu Leu

248

385 390 395 400

Glu Ala Asn Ser Val Phe Arg Gln Lys Val Glu Met Ser Tyr Gln Arg
 405 410 415

Met Glu Asn Pro Gly Cys His Val Val Asp Ala Ser Pro Ser Arg Glu
 420 425 430

Lys Val Leu Gln Thr Val Leu Ser Leu Ile Gln Asn Ser Phe Ser Glu
 435 440 445

Pro

<210> 206

<211> 590

<212> PRT

<213> Homo sapien

<400> 206

Pro Lys Ala Asn Glu Gln Leu Asn Arg Arg Ser Gln Arg Leu Gln Gln
 1 5 10 15

Leu Thr Glu Val Ser Arg Arg Ser Leu Arg Ser Arg Glu Ile Gln Gly
 20 25 30

Gln Val Gln Ala Val Lys Gln Ser Leu Pro Pro Thr Lys Lys Glu Gln
 35 40 45

Cys Ser Ser Thr Gln Ser Lys Ser Asn Lys Thr Ser Gln Lys His Val
 50 55 60

Lys Arg Lys Val Leu Glu Val Lys Ser Asp Ser Lys Glu Asp Glu Asn
 65 70 75 80

Leu Val Ile Asn Glu Val Ile Asn Ser Pro Lys Gly Lys Lys Arg Lys
 85 90 95

Val Glu His Gln Thr Ala Cys Ala Cys Ser Ser Gln Cys Met Gln Gly
 100 105 110

Ser Glu Lys Cys Pro Gln Lys Thr Thr Arg Arg Asp Glu Thr Lys Pro
 115 120 125

Val Pro Val Thr Ser Glu Val Lys Arg Ser Lys Met Ala Thr Ser Val
 130 135 140

Val Pro Lys Lys Asn Glu Met Lys Lys Ser Val His Thr Gln Val Asn
145 150 155 160

Thr Asn Thr Thr Leu Pro Lys Ser Pro Gln Pro Ser Val Pro Glu Gln
165 170 175

Ser Asp Asn Glu Leu Glu Gln Ala Gly Lys Ser Lys Arg Gly Ser Ile
180 185 190

Leu Gln Leu Cys Glu Glu Ile Ala Gly Glu Ile Glu Ser Asp Asn Val
195 200 205

Glu Val Lys Lys Glu Ser Ser Gln Met Glu Ser Val Lys Glu Glu Lys
210 215 220

Pro Thr Glu Ile Lys Leu Glu Glu Thr Ser Val Glu Arg Gln Ile Leu
225 230 235 240

His Gln Lys Glu Thr Asn Gln Asp Val Gln Cys Asn Arg Phe Phe Pro
245 250 255

Ser Arg Lys Thr Lys Pro Val Lys Cys Ile Leu Asn Gly Ile Asn Ser
260 265 270

Ser Ala Lys Lys Asn Ser Asn Trp Thr Lys Ile Lys Leu Ser Lys Phe
275 280 285

Asn Ser Val Gln His Asn Lys Leu Asp Ser Gln Val Ser Pro Lys Leu
290 295 300

Gly Leu Leu Arg Thr Ser Phe Ser Pro Pro Ala Leu Glu Met His His
305 310 315 320

Pro Val Thr Gln Ser Thr Phe Leu Gly Thr Lys Leu His Asp Arg Asn
325 330 335

Ile Thr Cys Gln Gln Glu Lys Met Lys Glu Ile Asn Ser Glu Glu Val
340 345 350

Lys Ile Asn Asp Ile Thr Val Glu Ile Asn Lys Thr Thr Glu Arg Ala
355 360 365

Pro Glu Asn Cys His Leu Ala Asn Glu Ile Lys Pro Ser Asp Pro Pro
370 375 380

250

Leu Asp Asn Gln Met Lys His Ser Phe Asp Ser Ala Ser Asn Lys Asn
385 390 395 400

Phe Ser Gln Cys Leu Glu Ser Lys Leu Glu Asn Ser Pro Val Glu Asn
405 410 415

Val Thr Ala Ala Ser Thr Leu Leu Ser Gln Ala Lys Ile Asp Thr Gly
420 425 430

Glu Asn Lys Phe Pro Gly Ser Ala Pro Gln Gln His Ser Ile Leu Ser
435 440 445

Asn Gln Thr Ser Lys Ser Ser Asp Asn Arg Glu Thr Pro Arg Asn His
450 455 460

Ser Leu Pro Lys Cys Asn Ser His Leu Glu Ile Thr Ile Pro Lys Asp
465 470 475 480

Leu Lys Leu Lys Glu Ala Glu Lys Thr Asp Glu Lys Gln Leu Ile Ile
485 490 495

Asp Ala Gly Gln Lys Arg Phe Gly Ala Val Ser Cys Asn Val Cys Gly
500 505 510

Met Leu Tyr Thr Ala Ser Asn Pro Glu Asp Glu Thr Gln His Leu Leu
515 520 525

Phe His Asn Gln Phe Ile Ser Ala Val Lys Tyr Val Val Leu Leu Ile
530 535 540

Asn His His Glu Cys Gly Ser Glu Glu Glu Phe Ile Thr Ser Leu Phe
545 550 555 560

Leu Ser Met Phe Asn Phe Arg Tyr Thr Gln Arg Ser Phe Ser Phe Pro
565 570 575

Ile Arg Phe Leu Glu Gly Leu Glu Glu Arg Lys Asn Ser Gly
580 585 590

<210> 207

<211> 661

<212> PRT

<213> Homo sapien

<400> 207

Met Gln Gly Ser Glu Lys Cys Pro Gln Lys Thr Thr Arg Arg Asp Glu
1 5 10 15

Thr Lys Pro Val Pro Val Thr Ser Glu Val Lys Arg Ser Lys Met Ala
20 25 30

Thr Ser Val Val Pro Lys Lys Asn Glu Met Lys Lys Ser Val His Thr
35 40 45

Gln Val Asn Thr Asn Thr Thr Leu Pro Lys Ser Pro Gln Pro Ser Val
50 55 60

Pro Glu Gln Ser Asp Asn Glu Leu Glu Gln Ala Gly Lys Ser Lys Arg
65 70 75 80

Gly Ser Ile Leu Gln Leu Cys Glu Glu Ile Ala Gly Glu Ile Glu Ser
85 90 95

Asp Asn Val Glu Val Lys Lys Glu Ser Ser Gln Met Glu Ser Val Lys
100 105 110

Glu Glu Lys Pro Thr Glu Ile Lys Leu Glu Glu Thr Ser Val Glu Arg
115 120 125

Gln Ile Leu His Gln Lys Glu Thr Asn Gln Asp Val Gln Cys Asn Arg
130 135 140

Phe Phe Pro Ser Arg Lys Thr Lys Pro Val Lys Cys Ile Leu Asn Gly
145 150 155 160

Ile Asn Ser Ser Ala Lys Lys Asn Ser Asn Trp Thr Lys Ile Lys Leu
165 170 175

Ser Lys Phe Asn Ser Val Gln His Asn Lys Leu Asp Ser Gln Val Ser
180 185 190

Pro Lys Leu Gly Leu Leu Arg Thr Ser Phe Ser Pro Pro Ala Leu Glu
195 200 205

Met His His Pro Val Thr Gln Ser Thr Phe Leu Gly Thr Lys Leu His
210 215 220

Asp Arg Asn Ile Thr Cys Gln Gln Glu Lys Met Lys Glu Ile Asn Ser
225 230 235 240

Glu Glu Val Lys Ile Asn Asp Ile Thr Val Glu Ile Asn Lys Thr Thr
245 250 255

252

Glu Arg Ala Pro Glu Asn Cys His Leu Ala Asn Glu Ile Lys Pro Ser
260 265 270

Asp Pro Pro Leu Asp Asn Gln Met Lys His Ser Phe Asp Ser Ala Ser
275 280 285

Asn Lys Asn Phe Ser Gln Cys Leu Glu Ser Lys Leu Glu Asn Ser Pro
290 295 300

Val Glu Asn Val Thr Ala Ala Ser Thr Leu Leu Ser Gln Ala Lys Ile
305 310 315 320

Asp Thr Gly Glu Asn Lys Phe Pro Gly Ser Ala Pro Gln Gln His Ser
325 330 335

Ile Leu Ser Asn Gln Thr Ser Lys Ser Ser Asp Asn Arg Glu Thr Pro
340 345 350

Arg Asn His Ser Leu Pro Lys Cys Asn Ser His Leu Glu Ile Thr Ile
355 360 365

Pro Lys Asp Leu Lys Leu Lys Glu Ala Glu Lys Thr Asp Glu Lys Gln
370 375 380

Leu Ile Ile Asp Ala Gly Gln Lys Arg Phe Gly Ala Val Ser Cys Asn
385 390 395 400

Val Cys Gly Met Leu Tyr Thr Ala Ser Asn Pro Glu Asp Glu Thr Gln
405 410 415

His Leu Leu Phe His Asn Gln Phe Ile Ser Ala Val Lys Tyr Val Val
420 425 430

Leu Leu Ile Asn His His Glu Cys Gly Ser Glu Glu Glu Phe Ile Thr
435 440 445

Ser Leu Phe Leu Ser Met Phe Asn Phe Arg Tyr Thr Gln Arg Ser Phe
450 455 460

Ser Phe Pro Ile Arg Phe Leu Glu Gly Trp Lys Lys Glu Arg Ile Leu
465 470 475 480

Ala Glu Tyr Pro Asp Gly Arg Ile Ile Met Val Leu Pro Glu Asp Pro
485 490 495

253

Lys Tyr Ala Leu Lys Lys Val Asp Glu Ile Arg Glu Met Val Asp Asn
 500 505 510

Asp Leu Gly Phe Gln Gln Ala Pro Leu Met Cys Tyr Ser Arg Thr Lys
 515 520 525

Thr Leu Leu Phe Ile Ser Asn Asp Lys Lys Val Val Gly Cys Leu Ile
 530 535 540

Ala Glu His Ile Gln Trp Gly Tyr Arg Val Ile Glu Glu Lys Leu Pro
 545 550 555 560

Val Ile Arg Ser Glu Glu Glu Lys Val Arg Phe Glu Arg Gln Lys Ala
 565 570 575

Trp Cys Cys Ser Thr Leu Pro Glu Pro Ala Ile Cys Gly Ile Ser Arg
 580 585 590

Ile Trp Val Phe Ser Met Met Arg Arg Lys Lys Ile Ala Ser Arg Met
 595 600 605

Ile Glu Cys Leu Arg Ser Asn Phe Ile Tyr Gly Ser Tyr Leu Ser Lys
 610 615 620

Glu Glu Ile Ala Phe Ser Asp Pro Thr Pro Asp Gly Lys Leu Phe Ala
 625 630 635 640

Thr Gln Tyr Cys Gly Thr Gly Gln Phe Leu Val Tyr Asn Phe Ile Asn
 645 650 655

Gly Gln Asn Ser Thr
 660

<210> 208
 <211> 157
 <212> PRT
 <213> Homo sapien

<400> 208

Met Thr Thr Val Glu Arg Gly Cys Gly Ser Gly Ala Ala Trp Arg Ala
 1 5 10 15

Val Gln Cys Arg Ala Gly Val Ser Gln Gly Leu Val Ala Thr Val Glu
 20 25 30

Arg Gly Cys Gly Ser Gly Ser Pro Ala Cys Ser Pro Val Pro Gly
 35 40 45

Arg Ser Leu Ala Glu Cys Ser Leu Thr Pro Pro Arg Gly Ser Pro Gly
50 55 60

Pro Tyr Arg Leu Pro Gln Leu Gln Ser Trp Val Pro Ser Asp Ala Val
65 70 75 80

Ala Gly Gln Arg Glu Ala Glu Ala Gly Ser Pro Arg Glu Ala Trp Ala
85 90 95

Pro Ser Pro Gly His Gly Cys Pro Ser Arg Ser Ser Ser Leu Gln Pro
100 105 110

Gln Ser Gln Gly Asp Val Gly Thr Gly Val Lys Ser Gly Trp Ser Val
115 120 125

Ala Leu Arg Pro Gln Glu Arg Tyr Gly Leu Lys Pro Ala Ala Arg Ala
130 135 140

Cys His Thr Arg Val Gly Pro Pro Leu His Ile Leu Arg
145 150 155

<210> 209

<211> 269

<212> PRT

<213> Homo sapien

<400> 209

Met Asp Arg Pro Pro Gly Gln Val Lys Ala Ala Thr Ser Asp Leu Glu
1 5 10 15

His Tyr Asp Lys Thr Arg His Glu Glu Phe Lys Lys Tyr Glu Met Met
20 25 30

Lys Glu His Glu Arg Arg Glu Tyr Leu Lys Thr Leu Asn Glu Glu Lys
35 40 45

Arg Lys Glu Glu Glu Ser Lys Phe Glu Glu Met Lys Lys Lys His Glu
50 55 60

Asn His Pro Lys Val Asn His Pro Gly Ser Lys Asp Gln Leu Lys Glu
65 70 75 80

Val Trp Glu Glu Thr Asp Gly Leu Asp Pro Asn Asp Phe Asp Pro Lys
85 90 95

255

Thr Phe Phe Lys Leu His Asp Val Asn Ser Asp Gly Phe Leu Asp Glu
100 105 110

Gln Glu Leu Glu Ala Leu Phe Thr Lys Glu Leu Glu Lys Val Tyr Asp
115 120 125

Pro Lys Asn Glu Glu Asp Asp Met Val Glu Met Glu Glu Glu Arg Leu
130 135 140

Arg Met Arg Glu His Val Met Asn Glu Val Asp Thr Asn Lys Asp Arg
145 150 155 160

Leu Val Thr Leu Glu Glu Phe Leu Lys Ala Thr Glu Lys Lys Glu Phe
165 170 175

Leu Glu Pro Asp Ser Trp Glu Thr Leu Asp Gln Gln Phe Phe Thr
180 185 190

Glu Glu Glu Leu Lys Glu Tyr Glu Asn Ile Ile Ala Leu Gln Glu Asn
195 200 205

Glu Leu Lys Lys Ala Asp Glu Leu Gln Lys Gln Lys Glu Glu Leu
210 215 220

Gln Arg Gln His Asp Gln Leu Glu Ala Gln Lys Leu Glu Tyr His Gln
225 230 235 240

Val Ile Gln Gln Met Glu Gln Lys Lys Leu Gln Gln Gly Ile Pro Pro
245 250 255

Ser Gly Pro Ala Gly Glu Leu Lys Phe Glu Pro His Ile
260 265

<210> 210
<211> 363
<212> PRT
<213> Homo sapien

<400> 210

Met Arg Trp Arg Thr Ile Leu Leu Gln Tyr Cys Phe Leu Leu Ile Thr
1 5 10 15

Cys Leu Leu Thr Ala Leu Glu Ala Val Pro Ile Asp Ile Asp Lys Thr
20 25 30

Lys Val Gln Asn Ile His Pro Val Glu Ser Ala Lys Ile Glu Pro Pro
35 40 45

Asp Thr Gly Leu Tyr Tyr Asp Glu Tyr Leu Lys Gln Val Ile Asp Val
50 55 60

Leu Glu Thr Asp Lys His Phe Arg Glu Lys Leu Gln Lys Ala Asp Ile
65 70 75 80

Glu Glu Ile Lys Ser Gly Arg Leu Ser Lys Glu Leu Asp Leu Val Ser
85 90 95

His His Val Arg Thr Lys Leu Asp Glu Leu Lys Arg Gln Glu Val Gly
100 105 110

Arg Leu Arg Met Leu Ile Lys Ala Lys Leu Asp Ser Leu Gln Asp Ile
115 120 125

Gly Met Asp His Gln Ala Leu Leu Lys Gln Phe Asp His Leu Asn His
130 135 140

Leu Asn Pro Asp Lys Phe Glu Ser Thr Asp Leu Asp Met Leu Ile Lys
145 150 155 160

Ala Ala Thr Ser Asp Leu Glu His Tyr Asp Lys Thr Arg His Glu Glu
165 170 175

Phe Lys Lys Tyr Glu Met Met Lys Glu His Glu Arg Arg Glu Tyr Leu
180 185 190

Lys Thr Leu Asn Glu Glu Lys Arg Lys Glu Glu Glu Ser Lys Phe Glu
195 200 205

Glu Met Lys Lys His Glu Asn His Pro Lys Val Asn His Pro Gly
210 215 220

Ser Lys Asp Gln Leu Lys Glu Val Trp Glu Glu Thr Asp Gly Leu Asp
225 230 235 240

Pro Asn Asp Phe Asp Pro Lys Thr Phe Phe Lys Leu His Asp Val Asn
245 250 255

Ser Asp Gly Phe Leu Asp Glu Gln Glu Leu Glu Ala Leu Phe Thr Lys
260 265 270

Glu Leu Glu Lys Val Tyr Asp Pro Lys Asn Glu Glu Asp Asp Met Val
275 280 285

257

Glu Met Glu Glu Glu Arg Leu Arg Met Arg Glu His Val Met Asn Glu
290 295 300

Val Asp Thr Asn Lys Asp Arg Leu Val Thr Leu Glu Glu Phe Leu Lys
305 310 315 320

Ala Thr Glu Lys Lys Glu Phe Leu Glu Pro Asp Ser Trp Glu Val Ile
325 330 335

Gln Gln Met Glu Gln Lys Lys Leu Gln Gln Gly Ile Pro Pro Ser Gly
340 345 350

Pro Ala Gly Glu Leu Lys Phe Glu Pro His Ile
355 360

<210> 211
<211> 420
<212> PRT
<213> Homo sapien

<400> 211

Met Arg Trp Arg Thr Ile Leu Leu Gln Tyr Cys Phe Leu Leu Ile Thr
1 5 10 15

Cys Leu Leu Thr Ala Leu Glu Ala Val Pro Ile Asp Ile Asp Lys Thr
20 25 30

Lys Val Gln Asn Ile His Pro Val Glu Ser Ala Lys Ile Glu Pro Pro
35 40 45

Asp Thr Gly Leu Tyr Tyr Asp Glu Tyr Leu Lys Gln Val Ile Asp Val
50 55 60

Leu Glu Thr Asp Lys His Phe Arg Glu Lys Leu Gln Lys Ala Asp Ile
65 70 75 80

Glu Glu Ile Lys Ser Gly Arg Leu Ser Lys Glu Leu Asp Leu Val Ser
85 90 95

His His Val Arg Thr Lys Leu Asp Glu Leu Lys Arg Gln Glu Val Gly
100 105 110

Arg Leu Arg Met Leu Ile Lys Ala Lys Leu Asp Ser Leu Gln Asp Ile
115 120 125

Gly Met Asp His Gln Ala Leu Leu Lys Gln Phe Asp His Leu Asn His

258

130

135

140

Leu Asn Pro Asp Lys Phe Glu Ser Thr Asp Leu Asp Met Leu Ile Lys
145 150 155 160

Ala Ala Thr Ser Asp Leu Glu His Tyr Asp Lys Thr Arg His Glu Glu
165 170 175

Phe Lys Lys Tyr Glu Met Met Lys Glu His Glu Arg Arg Glu Tyr Leu
180 185 190

Lys Thr Leu Asn Glu Glu Lys Arg Lys Glu Glu Glu Ser Lys Phe Glu
195 200 205

Glu Met Lys Lys His Glu Asn His Pro Lys Val Asn His Pro Gly
210 215 220

Ser Lys Asp Gln Leu Lys Glu Val Trp Glu Glu Thr Asp Gly Leu Asp
225 230 235 240

Pro Asn Asp Phe Asp Pro Lys Thr Phe Phe Lys Leu His Asp Val Asn
245 250 255

Ser Asp Gly Phe Leu Asp Glu Gln Glu Leu Glu Ala Leu Phe Thr Lys
260 265 270

Glu Leu Glu Lys Val Tyr Asp Pro Lys Asn Glu Glu Asp Asp Met Val
275 280 285

Glu Met Glu Glu Glu Arg Leu Arg Met Arg Glu His Val Met Asn Glu
290 295 300

Val Asp Thr Asn Lys Asp Arg Leu Val Thr Leu Glu Glu Phe Leu Lys
305 310 315 320

Ala Thr Glu Lys Lys Glu Phe Leu Glu Pro Asp Ser Trp Glu Thr Leu
325 330 335

Asp Gln Gln Gln Phe Phe Thr Glu Glu Glu Leu Lys Glu Tyr Glu Asn
340 345 350

Ile Ile Ala Leu Gln Glu Asn Glu Leu Lys Lys Ala Asp Glu Leu
355 360 365

Gln Lys Gln Lys Glu Glu Leu Gln Arg Gln His Asp Gln Leu Glu Ala
370 375 380

Gln Lys Leu Glu Tyr His Gln Val Ile Gln Gln Met Glu Gln Lys Lys
385 390 395 400

Leu Gln Gln Gly Ile Pro Pro Ser Gly Pro Ala Gly Glu Leu Lys Phe
405 410 415

Glu Pro His Ile
420

<210> 212
<211> 162
<212> PRT
<213> Homo sapien

<400> 212

Met Gln Thr Ser Val Thr Trp Glu Ile Pro Phe Pro Thr Asn Ser Leu
1 5 10 15

Val Val Lys Leu His Ser Met Asp Lys Ile Thr Tyr Tyr His Lys Ile
20 25 30

Lys Lys Cys Ile Phe Ser Ala Leu Arg Ala Arg Asn Thr Arg Arg Ser
35 40 45

Ile Lys Leu Asp Gly Lys Gly Glu Pro Lys Gly Ala Lys Arg Ala Lys
50 55 60

Pro Val Lys Tyr Thr Ala Ala Lys Leu His Glu Lys Gly Val Leu Leu
65 70 75 80

Asp Ile Asp Asp Leu Gln Thr Asn Gln Phe Lys Asn Val Thr Phe Asp
85 90 95

Ile Ile Ala Thr Glu Asp Val Gly Ile Phe Asp Val Arg Ser Lys Phe
100 105 110

Leu Gly Val Glu Met Glu Lys Val Gln Leu Asn Ile Gln Asp Leu Leu
115 120 125

Gln Met Gln Tyr Glu Gly Val Ala Val Met Lys Met Phe Asp Lys Val
130 135 140

Lys Val Asn Val Asn Leu Leu Ile Tyr Leu Leu Asn Lys Lys Phe Tyr
145 150 155 160

260

Gly Lys

<210> 213
<211> 69
<212> PRT
<213> Homo sapien

<400> 213

Tyr Phe Thr Leu Phe Tyr Tyr Lys Phe Arg Ser Leu Cys Phe Thr Ile
1 5 10 15

Asn Ser Asp Tyr Pro Asn Ile Phe Leu Ile Leu Cys Gly Asn Ala Asp
20 25 30

Phe Leu Leu Leu Arg Ser Gly Asn Ile Leu His Cys Leu His Ser Ser
35 40 45

His Gly Thr Trp Lys Phe Leu Lys Val Ile Tyr Asp Thr His Phe Leu
50 55 60

Cys Met Tyr Ser Asn
65

<210> 214
<211> 42
<212> PRT
<213> Homo sapien

<400> 214

Gln Ser Ser Ala Glu Ala Gly Gly Asp Glu Arg Glu Ile Asn Thr
1 5 10 15

Tyr Gly Arg Trp Ala Leu Met Gln Cys Glu Arg Arg Ser Val Met Asp
20 25 30

Val Arg Gly Arg Gly Thr Ser Glu Leu Pro
35 40

<210> 215
<211> 172
<212> PRT
<213> Homo sapien

<400> 215

Gly Thr Gly Leu Pro Trp His Ser Thr Pro Ala Gln Leu Ala Leu Ala
1 5 10 15

261

Gly Leu Arg Gln Ala Gln Pro His Pro Gln Gln Gln Arg Leu His Gln
20 25 30

Pro Gly Leu Arg Gly Val Asp Ala His Gly Ser Ala Ala His Val Pro
35 40 45

Gln Ala Val Pro Gln Ala Val Arg Ala His Pro Pro Gly Gln Leu Leu
50 55 60

Ser Trp Ala Ala Ala Val Cys Leu Leu Cys Gln His His Leu Gln Leu
65 70 75 80

Pro Gly Lys Lys Arg Asn Ser Thr Leu Tyr Ile Thr Met Leu Leu Ile
85 90 95

Val Pro Val Ile Val Ala Gly Ala Ile Ile Val Leu Leu Leu Tyr Leu
100 105 110

Lys Arg Leu Lys Ile Ile Ile Phe Pro Pro Pro Ile Pro Asp Pro Gly Lys
115 120 125

Ile Phe Lys Glu Met Phe Gly Asp Gln Asn Asp Asp Thr Leu His Trp
130 135 140

Lys Lys Tyr Asp Ile Tyr Glu Lys Gln Thr Lys Glu Glu Thr Asp Ser
145 150 155 160

Val Val Leu Ile Glu Asn Leu Lys Lys Ala Ser Gln
165 170

<210> 216
<211> 134
<212> PRT
<213> Homo sapien

<400> 216

Tyr Gly His Ile Arg Gln Gly Asn Tyr Ser Ala Gly Leu Pro Arg Cys
20 25 30

Val Tyr Cys Val Asn Ile Thr Tyr Asn Tyr Leu Gly Lys Lys Arg Asn
35 40 45

Ser Thr Leu Tyr Ile Thr Met Leu Leu Ile Val Pro Val Ile Val Ala
50 55 60

Gly Ala Ile Ile Val Leu Leu Leu Tyr Leu Lys Arg Leu Lys Ile Ile
65 70 75 80

Ile Phe Pro Pro Ile Pro Asp Pro Gly Lys Ile Phe Lys Glu Met Phe
85 90 95

Gly Asp Gln Asn Asp Asp Thr Leu His Trp Lys Lys Tyr Asp Ile Tyr
100 105 110

Glu Lys Gln Thr Lys Glu Glu Thr Asp Ser Val Val Leu Ile Glu Asn
115 120 125

Leu Lys Lys Ala Ser Gln
130

<210> 217

<211> 396

<212> PRT

<213> Homo sapien

<400> 217

Met Leu Met Ala Lys Gly Lys Leu Lys Pro Thr Gln Asn Ala Ser Glu
1 5 10 15

Lys Leu Gln Ala Pro Gly Lys Gly Leu Thr Ser Asn Lys Ser Lys Asp
20 25 30

Asp Leu Val Val Ala Glu Val Glu Ile Asn Asp Val Pro Leu Thr Cys
35 40 45

Arg Asn Leu Leu Thr Arg Gly Gln Thr Gln Asp Glu Ile Ser Arg Leu
50 55 60

Ser Gly Ala Ala Val Ser Thr Arg Gly Arg Phe Met Thr Thr Glu Glu
65 70 75 80

Lys Ala Lys Val Gly Pro Gly Asp Arg Pro Leu Tyr Leu His Val Gln
85 90 95

Gly Gln Thr Arg Glu Leu Val Asp Arg Ala Val Asn Arg Ile Lys Glu
100 105 110

Ile Ile Thr Asn Gly Val Val His Gln Pro Ala Pro Ile Ala Gln Leu
115 120 125

263

Ser Pro Ala Val Ser Gln Lys Pro Pro Phe Gln Ser Gly Met His Tyr
130 135 140

Val Gln Asp Lys Leu Phe Val Gly Leu Glu His Ala Val Pro Thr Phe
145 150 155 160

Asn Val Lys Glu Lys Val Glu Gly Pro Gly Cys Ser Tyr Leu Gln His
165 170 175

Ile Gln Ile Glu Thr Gly Ala Lys Val Phe Leu Arg Gly Lys Gly Ser
180 185 190

Gly Cys Ile Glu Pro Ala Ser Gly Arg Glu Ala Phe Glu Pro Met Tyr
195 200 205

Ile Tyr Ile Ser His Pro Lys Pro Glu Gly Leu Ala Ala Ala Lys Lys
210 215 220

Leu Cys Glu Asn Leu Leu Gln Thr Val His Ala Glu Tyr Ser Arg Phe
225 230 235 240

Val Asn Gln Ile Asn Thr Ala Val Pro Leu Pro Gly Tyr Thr Gln Pro
245 250 255

Ser Ala Ile Ser Ser Val Pro Pro Gln Pro Pro Tyr Tyr Pro Ser Asn
260 265 270

Gly Tyr Gln Ser Gly Tyr Pro Val Val Pro Pro Pro Gln Gln Pro Val
275 280 285

Gln Pro Pro Tyr Gly Val Pro Ser Ile Val Pro Pro Ala Val Ser Leu
290 295 300

Ala Pro Gly Val Leu Pro Ala Leu Pro Thr Gly Val Pro Pro Val Pro
305 310 315 320

Thr Gln Tyr Pro Ile Thr Gln Val Gln Pro Pro Ala Ser Thr Gly Gln
325 330 335

Ser Pro Met Gly Gly Pro Phe Ile Pro Ala Ala Pro Val Lys Thr Ala
340 345 350

Leu Pro Ala Gly Pro Gln Pro Gln Pro Gln Pro Gln Pro Pro Leu Pro
355 360 365

Ser Gln Pro Gln Ala Gln Lys Arg Arg Phe Thr Glu Glu Leu Pro Asp

264

370

375

380

Glu Arg Glu Ser Gly Leu Leu Gly Tyr Gln Val Lys
385 390 395

<210> 218
<211> 255
<212> PRT
<213> Homo sapien

<400> 218

Met His Tyr Val Gln Asp Lys Leu Phe Val Gly Leu Glu His Ala Val
1 5 10 15

Pro Thr Phe Asn Val Lys Glu Lys Val Glu Gly Pro Gly Cys Ser Tyr
20 25 30

Leu Gln His Ile Gln Ile Glu Thr Gly Ala Lys Val Phe Leu Arg Gly
35 40 45

Lys Gly Ser Gly Cys Ile Glu Pro Ala Ser Gly Arg Glu Ala Phe Glu
50 55 60

Pro Met Tyr Ile Tyr Ile Ser His Pro Lys Pro Glu Gly Leu Ala Ala
65 70 75 80

Ala Lys Lys Leu Cys Glu Asn Leu Leu Gln Thr Val His Ala Glu Tyr
85 90 95

Ser Arg Phe Val Asn Gln Ile Asn Thr Ala Val Pro Leu Pro Gly Tyr
100 105 110

Thr Gln Pro Ser Ala Ile Ser Ser Val Pro Pro Gln Pro Pro Tyr Tyr
115 120 125

Pro Ser Asn Gly Tyr Gln Ser Gly Tyr Pro Val Val Pro Pro Pro Gln
130 135 140

Gln Pro Val Gln Pro Pro Tyr Gly Val Pro Ser Ile Val Pro Pro Ala
145 150 155 160

Val Ser Leu Ala Pro Gly Val Leu Pro Ala Leu Pro Thr Gly Val Pro
165 170 175

Pro Val Pro Thr Gln Tyr Pro Ile Thr Gln Val Gln Pro Pro Ala Ser
180 185 190

265

Thr Gly Gln Ser Pro Met Gly Gly Pro Phe Ile Pro Ala Ala Pro Val
195 200 205

Lys Thr Ala Leu Pro Ala Gly Pro Gln Pro Gln Pro Gln Pro Gln Pro
210 215 220

Pro Leu Pro Ser Gln Pro Gln Ala Gln Lys Arg Arg Phe Thr Glu Glu
225 230 235 240

Leu Pro Asp Glu Arg Glu Ser Gly Leu Leu Gly Tyr Gln Val Lys
245 250 255

<210> 219

<211> 412

<212> PRT

<213> Homo sapien

<400> 219

Lys Ile Val Asp Val Ile Arg Gln Glu Val Leu Glu Ser Ser Gln Val
1 5 10 15

Thr Phe Val His His Leu Gln Ala Phe Ala Ser Lys Ile Thr Gly Met
20 25 30

Leu Leu Glu Leu Ser Pro Ala Gln Leu Leu Leu Leu Ala Ser Glu
35 40 45

Asp Ser Leu Arg Ala Arg Val Asp Glu Ala Met Glu Leu Ile Ile Ala
50 55 60

His Gly Arg Glu Asn Gly Ala Asp Ser Ile Leu Asp Leu Gly Leu Val
65 70 75 80

Asp Ser Ser Glu Lys Val Gln Gln Glu Asn Arg Lys Arg His Gly Ser
85 90 95

Ser Arg Ser Val Val Asp Met Asp Leu Asp Asp Thr Asp Asp Gly Asp
100 105 110

Asp Asn Ala Pro Leu Phe Tyr Gln Pro Gly Lys Arg Gly Phe Tyr Thr
115 120 125

Pro Arg Pro Gly Lys Asn Thr Glu Ala Arg Leu Asn Cys Phe Arg Asn
130 135 140

Ile Gly Arg Ile Leu Gly Leu Cys Leu Leu Gln Asn Glu Leu Cys Pro

266

145

150

155

160

Ile Thr Leu Asn Arg His Val Ile Lys Val Leu Leu Gly Arg Lys Val
165 170 175

Asn Trp His Asp Phe Ala Phe Phe Asp Pro Val Met Tyr Glu Ser Leu
180 185 190

Arg Gln Leu Ile Leu Ala Ser Gln Ser Ser Asp Ala Asp Ala Val Phe
195 200 205

Ser Ala Met Asp Leu Ala Phe Ala Ile Asp Leu Cys Lys Glu Glu Gly
210 215 220

Gly Gly Gln Val Glu Leu Ile Pro Asn Gly Val Asn Ile Pro Val Thr
225 230 235 240

Pro Gln Asn Val Tyr Glu Tyr Val Arg Lys Tyr Ala Glu His Arg Met
245 250 255

Leu Val Val Ala Glu Gln Pro Leu His Ala Met Arg Lys Gly Leu Leu
260 265 270

Asp Val Leu Pro Lys Asn Ser Leu Glu Asp Leu Thr Ala Glu Asp Phe
275 280 285

Arg Leu Leu Val Asn Gly Cys Gly Glu Val Asn Val Gln Met Leu Ile
290 295 300

Ser Phe Thr Ser Phe Asn Asp Glu Ser Gly Glu Asn Ala Glu Lys Leu
305 310 315 320

Leu Gln Phe Lys Arg Trp Phe Trp Ser Ile Val Glu Lys Met Ser Met
325 330 335

Thr Glu Arg Gln Asp Leu Val Tyr Phe Trp Thr Ser Ser Pro Ser Leu
340 345 350

Pro Ala Ser Glu Glu Gly Phe Gln Pro Met Pro Ser Ile Thr Ile Arg
355 360 365

Pro Pro Asp Asp Gln His Leu Pro Thr Ala Asn Thr Cys Ile Ser Arg
370 375 380

Leu Tyr Val Pro Leu Tyr Ser Ser Lys Gln Ile Leu Lys Gln Lys Leu
385 390 395 400

Leu Leu Ala Ile Lys Thr Lys Asn Phe Gly Phe Val
405 410

<210> 220
<211> 56
<212> PRT
<213> Homo sapien

<400> 220

Gly Lys Lys Lys Phe Asn Phe Gly Arg Leu Cys Tyr Leu Glu Ser Leu
1 5 10 15

Lys Phe Ser Ile Val Lys Met Asp Cys Ile Leu Leu Leu Thr Lys Ile
20 25 30

Ser Arg Ile Met Cys Gly Leu Leu Ile Ser Gly Met Leu Arg Ser Tyr
35 40 45

Ser Leu Thr Ile Lys Ile Leu Asn
50 55

<210> 221
<211> 430
<212> PRT
<213> Homo sapien

<400> 221

Glu Cys Pro Gly Arg Arg Asp Pro Gly Arg Gly Glu Arg Glu Gln Ser
1 5 10 15

Gly Val Arg Ala Ser Leu Trp Ala Gly Leu Gly Leu Gly Arg Arg
20 25 30

Cys Gly Leu Gly Arg Phe Gly Arg Gly Gly Arg Met Met Gly Arg
35 40 45

Val Arg Thr Leu Ala Gly Glu Cys Ser Ala Gln Ala Gln Ala Gln Ser
50 55 60

Leu Leu Ala Val Val Leu Ser Ala Pro Pro Ser Gly Gly Thr Pro Ser
65 70 75 80

Ala Arg Leu Ser Val Arg Ser Pro Ser Pro Arg Asp Pro Trp Gly Leu
85 90 95

Trp Ala Pro Val Leu Gln Met Thr Gly Ser Asn Glu Phe Lys Leu Asn

268

100

105

110

Gln Pro Pro Glu Asp Gly Ile Ser Ser Val Lys Phe Ser Pro Asn Thr
115 120 125

Ser Gln Phe Leu Leu Val Ser Ser Trp Asp Thr Ser Val Arg Leu Tyr
130 135 140

Asp Val Pro Ala Asn Ser Met Arg Leu Lys Tyr Gln His Thr Gly Ala
145 150 155 160

Val Leu Asp Cys Ala Phe Tyr Asp Pro Thr His Ala Trp Ser Gly Gly
165 170 175

Leu Asp His Gln Leu Lys Met His Asp Leu Asn Thr Asp Gln Glu Asn
180 185 190

Leu Val Gly Thr His Asp Ala Pro Ile Arg Cys Val Glu Tyr Cys Pro
195 200 205

Glu Val Asn Val Met Val Thr Gly Ser Trp Asp Gln Thr Val Lys Leu
210 215 220

Trp Asp Pro Arg Thr Pro Cys Asn Ala Gly Thr Phe Ser Gln Pro Glu
225 230 235 240

Lys Val Tyr Thr Leu Ser Val Ser Gly Asp Arg Leu Ile Val Gly Thr
245 250 255

Ala Gly Arg Arg Val Leu Val Trp Asp Leu Arg Asn Met Gly Tyr Val
260 265 270

Gln Gln Arg Arg Glu Ser Ser Leu Lys Tyr Gln Thr Arg Cys Ile Arg
275 280 285

Ala Phe Pro Asn Lys Gln Gly Tyr Val Leu Ser Ser Ile Glu Gly Arg
290 295 300

Val Ala Val Glu Tyr Leu Asp Pro Ser Pro Glu Val Gln Lys Lys Lys
305 310 315 320

Tyr Ala Phe Lys Cys His Arg Leu Lys Glu Asn Asn Ile Glu Gln Ile
325 330 335

Tyr Pro Val Asn Ala Ile Ser Phe His Asn Ile His Asn Thr Phe Ala
340 345 350

Thr Gly Gly Ser Asp Gly Phe Val Asn Ile Trp Asp Pro Phe Asn Lys
355 360 365

Lys Arg Leu Cys Gln Phe His Arg Tyr Pro Thr Ser Ile Ala Ser Leu
370 375 380

Ala Phe Ser Asn Asp Gly Thr Thr Leu Ala Ile Ala Ser Ser Tyr Met
385 390 395 400

Tyr Glu Met Asp Asp Thr Glu His Pro Glu Asp Gly Ile Phe Ile Arg
405 410 415

Gln Val Thr Asp Ala Glu Thr Lys Pro Lys Ser Pro Cys Thr
 420 425 430

<210> 222
<211> 385
<212> PRT
<213> *Homo sapien*

<400> 222

Met Gly Arg Val Arg Thr Leu Ala Gly Glu Cys Ser Ala Gln Ala Gln
1 5 10 15

Ala Gln Ser Leu Leu Ala Val Val Leu Ser Ala Pro Pro Ser Gly Gly
20 25 30

Thr Pro Ser Ala Arg Leu Ser Val Arg Ser Pro Ser Pro Arg Asp Pro
35 40 45

Trp Gly Leu Trp Ala Pro Val Leu Gln Met Thr Gly Ser Asn Glu Phe
50 55 60

Lys Leu Asn Gln Pro Pro Glu Asp Gly Ile Ser Ser Val Lys Phe Ser
65 70 75 80

Pro Asn Thr Ser Gln Phe Leu Leu Val Ser Ser Trp Asp Thr Ser Val
85 86 87 88 89 90 91 92 93 94 95

Arg Leu Tyr Asp Val Pro Ala Asn Ser Met Arg Leu Lys Tyr Gln His
100 105 110

Thr Gly Ala Val Leu Asp Cys Ala Phe Tyr Asp Pro Thr His Ala Trp
115 120 125

270

Ser Gly Gly Leu Asp His Gln Leu Lys Met His Asp Leu Asn Thr Asp
130 135 140

Gln Glu Asn Leu Val Gly Thr His Asp Ala Pro Ile Arg Cys Val Glu
145 150 155 160

Tyr Cys Pro Glu Val Asn Val Met Val Thr Gly Ser Trp Asp Gln Thr
165 170 175

Val Lys Leu Trp Asp Pro Arg Thr Pro Cys Asn Ala Gly Thr Phe Ser
180 185 190

Gln Pro Glu Lys Val Tyr Thr Leu Ser Val Ser Gly Asp Arg Leu Ile
195 200 205

Val Gly Thr Ala Gly Arg Arg Val Leu Val Trp Asp Leu Arg Asn Met
210 215 220

Gly Tyr Val Gln Gln Arg Arg Glu Ser Ser Leu Lys Tyr Gln Thr Arg
225 230 235 240

Cys Ile Arg Ala Phe Pro Asn Lys Gln Gly Tyr Val Leu Ser Ser Ile
245 250 255

Glu Gly Arg Val Ala Val Glu Tyr Leu Asp Pro Ser Pro Glu Val Gln
260 265 270

Lys Lys Lys Tyr Ala Phe Lys Cys His Arg Leu Lys Glu Asn Asn Ile
275 280 285

Glu Gln Ile Tyr Pro Val Asn Ala Ile Ser Phe His Asn Ile His Asn
290 295 300

Thr Phe Ala Thr Gly Ser Asp Gly Phe Val Asn Ile Trp Asp Pro
305 310 315 320

Phe Asn Lys Lys Arg Leu Cys Gln Phe His Arg Tyr Pro Thr Ser Ile
325 330 335

Ala Ser Leu Ala Phe Ser Asn Asp Gly Thr Thr Leu Ala Ile Ala Ser
340 345 350

Ser Tyr Met Tyr Glu Met Asp Asp Thr Glu His Pro Glu Asp Gly Ile
355 360 365

Phe Ile Arg Gln Val Thr Asp Ala Glu Thr Lys Pro Lys Ser Pro Cys

271

370

375

380

Thr

385

<210> 223

<211> 123

<212> PRT

<213> Homo sapien

<400> 223

Met	Pro	Ser	Ala	Met	Thr	Val	Tyr	Ala	Leu	Val	Val	Val	Ser	Tyr	Phe	
1						5								10		15

Leu	Ile	Thr	Gly	Gly	Ile	Ile	Tyr	Asp	Val	Ile	Val	Glu	Pro	Pro	Ser	
						20					25				30	

Val	Gly	Ser	Met	Thr	Asp	Glu	His	Gly	His	Gln	Arg	Pro	Val	Ala	Phe
						35				40				45	

Leu	Ala	Tyr	Arg	Val	Asn	Gly	Gln	Tyr	Ile	Met	Glu	Gly	Leu	Ala	Ser
						50			55				60		

Ser	Phe	Leu	Phe	Thr	Met	Gly	Gly	Leu	Gly	Phe	Ile	Ile	Leu	Asp	Arg
65					70					75				80	

Ser	Asn	Ala	Pro	Asn	Ile	Pro	Lys	Leu	Asn	Arg	Phe	Leu	Leu	Phe
						85			90				95	

Ile	Gly	Phe	Val	Cys	Val	Leu	Leu	Ser	Phe	Phe	Met	Ala	Arg	Val	Phe
						100			105				110		

Met	Arg	Met	Lys	Leu	Pro	Gly	Tyr	Leu	Met	Gly	
						115			120		

<210>	224
<211>	211
<212>	PRT
<213>	Homo sapien

<400> 224

Asn	Ile	Tyr	Leu	Leu	Ile	Leu	Lys	Cys	Phe	Lys	Lys	Ile	Lys	Lys
1					5					10			15	

Lys	Lys	Gln	Lys	Lys	Arg	Arg	Ala	Arg	Ala	Lys	Pro	Ala	Trp
							20			25			30

272

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 225
<211> 178
<212> PRT
<213> Homo sapien

<400> 225

Met Ala Arg Arg Pro Ala Gly Arg Glu Asn Ser Gly Val Pro Arg Gly
1 5 10 15

Leu Pro Lys Phe Ser Pro Pro Thr Phe Ser Ala Ala Thr Asn Val Arg
20 25 30

Ala Ala Gln Arg Gly Arg Pro Arg Arg Ala Pro Asp Ala Thr Arg Arg
35 40 45

Thr Ala Arg Ala Gly Thr Thr Pro Pro Arg His Gly Gln Pro Pro Ala
50 55 60

His Ala Arg Ala Ala Pro Ala His Asn Arg Gly Leu Pro Ser Cys Cys
65 70 75 80

Ser Arg Cys Arg Ala Lys Ala Arg Tyr Ala Arg Pro Arg Arg Ala Glu
85 90 95

Ala Ala Ala Arg Ala Arg Arg Ala Thr Pro Ala Ala Pro Gly Trp Arg
100 105 110

Gly Gly Gly Thr Ala Thr Arg Pro Thr Arg Arg Arg Ala Gly Thr Asn
115 120 125

Ala His Asp Pro His Arg Asn Gly Glu Gln Arg Pro Ser Gly Gln Arg
130 135 140

Arg Pro Arg Arg Gly Ser Arg Arg Arg Arg His Glu Thr Arg Glu Ser
145 150 155 160

Glu Arg Pro Leu Arg Gly Ala Gly Pro Gly Val Pro Gly Pro Thr Arg
165 170 175

Gly Gly

<210> 226
<211> 211
<212> PRT
<213> Homo sapien

<400> 226

Asn Ile Tyr Leu Leu Ile Leu Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

274

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 227
<211> 211
<212> PRT
<213> Homo sapien

<400> 227

Asn Ile Tyr Leu Leu Ile Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 228
<211> 211
<212> PRT
<213> Homo sapien

<400> 228

Asn Ile Tyr Leu Leu Ile Leu Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

276

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 229
<211> 211
<212> PRT
<213> Homo sapien

<400> 229

Asn Ile Tyr Leu Leu Ile Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 230
<211> 211
<212> PRT
<213> Homo sapien

<400> 230

Asn Ile Tyr Leu Leu Ile Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

278

Lys Lys Gln Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 231
<211> 211
<212> PRT
<213> Homo sapien

<400> 231

Asn Ile Tyr Leu Leu Ile Leu Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 232
<211> 211
<212> PRT
<213> Homo sapien

<400> 232

280

Asn Ile Tyr Leu Leu Ile Leu Lys Cys Phe Lys Lys Ile Lys Lys
1 5 10 15

Lys Lys Gln Lys Lys Arg Arg Ala Arg Arg Ala Lys Pro Ala Trp
20 25 30

Pro Trp Arg Gly Asp Pro Arg Gly Ala Lys Thr Val Ala Tyr Leu Ala
35 40 45

Ala Ser Pro Asn Ser Pro His Pro Pro Leu Ala Gln Arg Pro Thr Cys
50 55 60

Ala Pro Arg Ser Gly Gly Arg Asp Glu Arg Arg Thr Leu Arg Asp
65 70 75 80

Gly Arg Arg Gly Pro Ala Pro Arg His His Val Thr Gly Ser Arg Gln
85 90 95

Arg Thr Pro Gly Arg Arg Leu Leu Thr Thr Glu Val Cys Leu Val Ala
100 105 110

Ala Pro Gly Ala Glu Pro Arg Pro Ala Thr His Ala His Ala Gly Leu
115 120 125

Arg Gln Arg His Ala Arg Gly Val Gln Arg Arg Arg His Pro Ala Gly
130 135 140

Gly Gly Glu Ala Pro Gln His Gly Arg Arg Gly Glu Glu Arg Glu Gln
145 150 155 160

Thr His Thr His Thr Ala Thr Val Ser Asn Asp Arg Ala Ala Ser
165 170 175

Gly Asp Arg Gly Val Ala Ala Gly Asp Asp Ala Thr Arg Arg Ala Arg
180 185 190

Ala Arg Asp His Ser Glu Ala Pro Ala Arg Val Cys Gln Ala Arg Arg
195 200 205

Val Val Ala
210

<210> 233

<211> 24

<212> DNA

<213> Artificial sequence

281

<220>
<223> Synthetic

<400> 233
tggttgagaa gacatgaaaa tcca

24

<210> 234
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 234
aattccaccc tgtcaaccta aaaaaa

25

<210> 235
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 235
tgatttttgtt gtttccgaat ttcaaggcaa

29

<210> 236
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 236
agggggatta caatgatgga cc

22

<210> 237
<211> 18
<212> DNA
<213> Artificial sequence

<220>
<223> Synthetic

<400> 237
ttgcccaaggt gcgagctt

18

<210> 238
<211> 23
<212> DNA
<213> Artificial sequence

<220>

282

<223> Synthetic

<400> 238

agtgagcgct tagatggcca gca

23

<210> 239

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Synthetic

<400> 239

acaataaaatc agtaagcggtt ccagaa

26

<210> 240

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> Synthetic

<400> 240

caatctacat taaaaacata cacgtgaaca

30

<210> 241

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Synthetic

<400> 241

cttcttcacc tcctgagcca ctca

24

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

FADED TEXT OR DRAWING

BLURRED OR ILLEGIBLE TEXT OR DRAWING

SKEWED/SLANTED IMAGES

COLOR OR BLACK AND WHITE PHOTOGRAPHS

GRAY SCALE DOCUMENTS

LINES OR MARKS ON ORIGINAL DOCUMENT

REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.